lacittadinanzaeuropea online
Rivista di studi e documentazione sull'integrazione europea
ISSN 2785-4183
lceonline 3/2024 (*www.lceonline.eu*)

Comparative remarks on the regulation of predictive justice

Katia DE BLASIO*

SUMMARY: 1. An attempt to define predictive justice **2.** Shortcomings of algorithmic justice **3.** Ethical principles for the use of AI in judicial systems **4.** A "hard law" approach to predictive justice: the examples of the European Union, France, and Italy **5.** Guidelines, education, and corporations: the approach of the United States **6.** Final remarks

1. An attempt to define predictive justice

The idea of legal certainty goes back a long way, it was one of the core principles first of the French Revolution and then enshrined in Napoleon's codification. It is the means by which legal culture has pursued the idea of a clear and predictable law[1], attributes considered the necessary grounds for business to flourish in a bourgeois and then capitalistic society[2].

However, the advent of new technologies sheds new light on this principle and in particular on the prediction of judicial decisions. In broad terms, the expression "predictive justice", although not easy to define, indicates the use of technology to project into the future a statistic derived from past decisions that are similar to the case in question[3].

Lately, predictive justice has been gaining more and more popularity because of the increasing availability of artificial intelligence (AI) software which enables legal experts to calculate, for example, the chances of winning a court case, the amount of money that a person could obtain in damages and the arguments that could have a positive impact on the judge during the pleading[4]. AI programs could also be used to predict a future behaviour, for example the risk of recidivism during the trial or the pre- and post-trial phases[5].

---

* PhD Candidate in "Law and Innovation", University of Macerata.

[1] P. Grossi, *Storicità versus prevedibilità: sui caratteri di un diritto post-moderno,* in *Questione Giustizia,* 4/2018, p. 17. As was also clearly stated in the words of Holmes: "a legal duty so called is nothing but a prediction that if a man does or omits certain things he will be made to suffer in this or that way by judgment of the court; and so of a legal right", see O. W. Holmes, *The Path of the Law*, in *Harvard Law Review*, 10/1897, p. 458. Of course, we must distinguish the idea of legal certainty as it is understood in civil law systems as well as in common law systems, like the English one. In the former, the "positivistic" idea of certainty is linked to the content of written statutes, while in the latter, legal certainty is not to be associated with a static, authoritative text, but with the law applicable to a concrete case as it derives from the doctrine of binding legal precedent, see on this point L. Moccia, *Comparazione giuridica e diritto europeo,* Milano, 2005, p. 443.

[2] S. Benvenuti, S. Zolea, *European Courts And Predictive Justice: A Feasible Symbiosis?,* in *Opinio Juris in Comparatione,* 1/2023, pp. 56-57.

[3] L. Larret-Chahine, *La Justice Prédictive*, in E. Calzolaio (ed), *La decisione nel prisma dell'intelligenza artificiale*, Padova, 2020, p. 163. Websites collecting case law have been flourishing lately, as they enable lawyers and legal experts in general to know what the general interpretation of a certain rule is, on this topic see E. Gabellini, *Algoritmi decisionali e processo civile: limiti e prospettive.,* in *Rivista Trimestrale di Diritto e Procedura Civile,* 1/2022, pp. 64-67.

[4] L. Larret-Chahine*, La Justice Prédictive,* cit., p. 163.

[5] *Ex multis:* G. Padua, *Intelligenza artificiale e giudizio penale: scenari, limiti e prospettive*, in *Processo Penale e Giustizia*, 6/2021, p. 1498; D. L. Chen, *Judicial analytics and the great transformation of American Law,* in *Artificial Intelligence and Law* 27/2019, pp. 15-42; S. Fazel et al, *The predictive performance of criminal risk assessment tools used at sentencing: Systematic review of validation studies,* in *Journal of Criminal Justice,* 81/2022, pp. 1-9; E. Chelioudakis, *Risk Assessment*

The use of these predictive tools is widespread among law firms, especially the big ones based in the U.S., as AI helps them deliver their legal services in new, more cost-effective ways. In fact, in the last few years, especially in the wake of the 2008 global recession, clients are less willing to accept the indeterminacy of the hourly billing model prevalent in the U.S. and Canada[6] and are demanding faster and cheaper answers, which can be achieved through the use of technologies such as AI.

Moreover, not only in the U.S. is the practice of law becoming more complex over time but also in other legal systems. There is an increasing number of court decisions and new pieces of legislation coming from different national and supranational legislators[7] that pose new challenges for the justice system and lead to the belief that AI could be used to help lawyers quickly and effectively assess the merits of a case by accurately predicting the likely outcome of a particular litigation strategy[8].

As a result, big international law firms are investing heavily in AI experts[9] and developing their own AI tools[10] in order to gain access to personalized features that will make them competitive on the market[11]. This, in turn, is leading to changes in the internal organization of law firms so that the armies of young lawyers until now hired to perform boring, repetitive tasks will no longer be needed with the advent of AI. Also, the impact of these tools on legal education cannot be ignored, as future lawyers will be required to be well-skilled in the use of this technology.

In addition to the internal organization of law firms and legal education, the introduction of AI tools is also likely to have an impact on judicial activity. In fact, it is expected that these programs will be increasingly used by judges, who, especially in the United States, already have at their disposal some AI tools offered by the AI companies themselves[12].

As mentioned above, predictive justice programs are used not only to assess the merits of a case, but also to calculate recidivism rates, and particularly in the United States judges use these tools (defined as "risk assessment tools") in different phases of the criminal process[13], such as parole, sentencing, pretrial, and probation. The use of a risk assessment tool called COMPAS in the sentencing phase has generated a lot of discussion on the implementation of such technologies in the judiciary, as we will see in the following pages.

Since COMPAS is a proprietary software, whose algorithm is not disclosed, the U.S. Government itself has developed another tool called PATTERN to build more trust in risk assessment tools[14]. It is designed to predict the likelihood of general and violent recidivism three years after release and is based on static (e.g., criminal history) and dynamic factors (e.g., participation in education or drug treatment).

Tools in Criminal Justice: Is There a Need for Such Tools in Europe and Would Their Use Comply with European Data Protection Law?, in *Australian National University Journal of Law &Technology,* 2/2020, pp. 72-96; B. Custers, *AI in Criminal Law: An Overview of AI Applications in Substantive and Procedural Criminal Law,* in B.H.M. Custers & E. Fosch Villaronga (eds.), *Law and Artificial Intelligence,* Heidelberg, pp. 205-223; B. Garrett, C. Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice,* in *Duke Law School Public Law & Legal Theory Series*, 3/2023, pp. 1-52. In addition to prediction, AI is used for many other many tasks that were previously performed only by lawyers. Examples include legal research and e-discovery, which means that AI can be used to identify relevant elements in a large number of electronic documents; document automation, such as contract drafting, and case management.

[6] B. Alarie, A. Niblett, A. H. Yoon, *How Artificial Intelligence Will Affect the Practice of Law*, 7th Novembre 2017, p. 4.

[7] *Ibid.*

[8] *Ibid,* p. 10

[9] J. Henry, *Big Law's AI Talent War Aims to Influence Software Development,* 14 May 2024.

[10] For more information on the many AI tools implemented in American (or transnational) big law firms, check the law news website: https://www.lawfuel.com/the-biglaw-firms-using-ai-legaltech-to-redefine-legal-practices/, which also notes that DLA Piper, an international law firm with offices in many countries, has developed its own predictive tool.

[11] A. E. Davis, *The future of Law Firms (and Lawyers) in the Age of Artificial Intelligence,* in *Direito GV Law Review,* 1/2020, p. 2

[12] *U.S. Judiciary Receives Free Access to A.I. Legal Research Technology Through New Partnership Between American Bankruptcy Institute & Casetext,* 24th February 2020.

[13] For an overview of the risk assessment tools implemented in different jurisdictions of the United States, see: https://criminaljustice.tooltrack.org/tools, a database curated by the Berkman Klein Center at Harvard University.

[14] For more info on PATTERN: U.S. Department of Justice, *2022 Review and Revalidation of the First Step Act Risk Assessment Tool*, March 2023.

In Canada, a country that shares some similarities with the United States, risk assessment tools are used at various stages of the criminal process, but not without issues, which can be understood by looking at a famous case, *Ewert v. Canada*[15], where a risk assessment tool was used by the Correctional Service Canada (CSC), the government agency responsible for administering court-imposed sentencing of two years and more. Here, in a dispute over the accuracy of the data fed to the algorithm, the Court ruled that the CSC must "take all reasonable steps to ensure that any information about an offender that it uses is as accurate, up to date and complete as possible".

Predictive justice tools are also widespread in Europe: an example is Prédictice[16], a piece of software developed by a French company and able to predict the outcome of a trial based on the place where the lawsuit was filed.

Although a little late if compared to other countries, Italy too has been showing interest in the implementation of AI in the justice system. There are several ongoing projects[17] that have grown out of the collaboration between universities and courtrooms, one of which is called "Giustizia predittiva" run by the University of Venice, the Court of Appeal of Venice and the private company Deloitte. Within this project, researchers developed a replicable AI tool presented to the public in November 2022[18]. The algorithm has been trained on decisions related to labour and commercial law[19] and, similarly to Prédictice, is able to give answers to questions posed in a natural language (e.g. "is it possible to fire an employee because of her unjustified absences?")[20].

However, although these projects are very interesting, they do not include government bodies that could influence the policy on predictive justice. Therefore, it is worth mentioning another project that aims to develop a predictive tool for fiscal claims and involves the Italian Ministry of Economy and Finance and the Presidential Council of Fiscal Justice. This project is part of the so-called "Pro.di.gi.t"[21] (a broader project aiming at digitalizing fiscal justice). At the current stage, AI is being employed to summarize judgments, which will probably be used to train the predictive software[22], although the National Association of Italian Lawyers has already expressed some doubts about the tool, due to the lack of involvement of lawyers in its development[23].

All the examples analysed above show different implementations of AI as a support tool in the hands of judges and other subjects involved in pre- and post-trial phases. There are also examples

---

[15] In this case an inmate of indigenous descent claimed that the use of psychological and actuarial assessment tools by the Correctional Service of Canada (CSC) violated s. 24(1) of the *Corrections and Conditional Release Act* (CCRA), requiring CSC to "take all reasonable steps to ensure that any information about an offender that it uses is accurate, up to date and complete as possible", as well as ss. 7 and 15 of the *Charter.* The Court allowed the appeal in part, with two judges dissenting. In brief, the Court held that the CSC had failed to meet its obligations under s. 24(1) of the CCRA, because it had long been aware of concerns regarding cultural bias embedded in such tools, but it took no action to address them. With respect to the claims concerning ss. 7 and 15 of the *Charter,* the Court found that the claimant had failed to demonstrate that the tool was likely to operate differently in case of Indigenous offenders. Nor was it was established that the use of the tool resulted in an overestimation of the risk posed by Indigenous inmates, in harsher conditions of incarceration or in the denial of rehabilitative opportunities.

[16] To see more about Prédictice, https://predictice.com/fr.

[17] For example one project is being carried out by the Court of Appeal of Bari, https://ca-bari.giustizia.it/it/progetto_prevedibilita.page, another, called "Predictive Jurisprudence" is carried out by the Scuola Superiore Sant'Anna and the Courts of Pisa and Genova, https://www.predictivejurisprudence.eu/, another involves the Court of Cassation and the Scuola Universitaria Superiore IUSS of Pavia, https://www.iusspavia.it/sites/default/files/2022-08/c.s.%20Accordo%20 quadro%20 Cassazione-IUSS%20Pavia%20in%20tema%20di%20ricerca%20su%20materiale%20giuridico%20digitale. pdf.

[18] For more information on the outcome of this project, G. Musella, *"Giurisprudenza predittiva". Risultati operative e prospettive future,* in *Studi e ricerche Ca' Foscari,* 34/2023, pp.- 280-292.

[19] *Ibid,* p. 280.

[20] *Ibid,* p. 287.

[21] For more information about the project, check the website of the Italian Ministry of Economics and Finance: https://www.finanze.gov.it/it/Progetti-europei/PRO.DI.GI.T/

[22] L. Leoni, *Dalla (ir)reperibilità delle sentenze all'iperbole della c.d. giustizia predittiva,* in *Il Fisco,* 21/2024, p. 1989 ff.

[23] M. Masi, *Intervento della Presidente del Consiglio Nazionale Forense alla Cerimonia di inaugurazione dell'anno giudiziario tributario*, 14 March 2023.

of AI "making its own decisions", although great care must be taken not to sensationalize news that in reality is not as significant as it seems.

For example, some time ago, the news that a robot judge was being developed in Estonia went viral[24]. The report claimed that the Estonian Ministry of Justice was creating a "robot judge" to adjudicate small claims disputes of less than €7,000 in order to clear a backlog of cases for judges and court clerks. This news was subsequently denied by the Ministry[25], which called the article "misleading" and affirmed that projects of this sort had never been pursued in the Estonian public sector, and the only ongoing plans involving the use of AI were for the purpose of transcribing court hearings and anonymizing decisions.

In China, which is a highly technological country, AI has been utilized in many jurisdictions. It was reported that an AI software with the appearance of a woman had been implemented in a criminal court in Beijing to replace judges in repetitive basic work[26]. However, a closer reading of the news, revealed that AI was used to prioritize some cases before the court and not for the decision-making itself[27]. More interesting, in terms of the development of the so-called "robot-justice", is the experiment of the "Internet Courts", which were first set up in 2017 initially in Hangzhou and then in Beijing and Guangzhou to resolve online disputes (they have jurisdiction over Internet- related civil and administrative cases)[28]. These courts use AI for various purposes: for speech recognition and thus for transcribing court proceedings; for creating online decision documents and for automatic decision making. Moreover, the image of a human judge has been developed in these courts to identify the key words of the claims and provide appropriate responses[29].

As we can see from the cases cited above, AI is being used in many systems for predictive purposes, but to understand its potential and challenges, it is important to understand the basic principles of how it functions.

To put it simply, in AI programs data are codified (which means that they are transposed into a mathematical language understandable by the software) and then analysed and put in correlation according to the instructions given by the algorithm to obtain a final output[30]. Artificial intelligence programs do not have a creative intelligence[31]; they are not able to learn in the way that humans do; they simply create an output from the input (data) processed. Therefore, it is important to focus on the quality of the data set that is fed to the algorithm, otherwise one could have a "garbage in-garbage out" effect.

In this article we will not deal with robot-justice, which seems to be employed only in China, but will focus on tools like Prédictice or COMPAS, which could be referenced as Decision Support Systems (DSS)[32], conceived to be support devices in the hands of legal experts.

Such tools may seem harmless, but they are not, especially if used in the decision-making process, because it is difficult to state to what extent a judge can delegate a decision or some parts of it to a piece of software. It has been shown, in fact, that humans tend to rely on the outcome of highly technological and independent tools[33], and this could lead to severe negative consequences,

---

[24] E. Niler, *Can AI Be a Fair Judge in Court? Estonia Thinks So*, 29 March 2019.
[25] Republic of Estonia, Ministry of Justice, *Estonia does not develop AI Judge*, 16 February 2022.
[26] For further information on this experiment, check the website of the Chinese Government: https://english.bjinternetcourt. gov.cn/2019-07/01/c_190.htm.
[27] A. Santosuosso, G. Sartor, *La giustizia predittiva: una visione realistica*, in *Giurisprudenza italiana*, 7/2022, p. 1760.
[28] D. Clementi, C. Comberiati, *Digital justice as a tool of socio-juridical control: the cases of the United States of America and the People's Republic of China,* in *La cittadinanza europea online,* 1/2023, pp. 28-29.
[29] *Ibid,* p. 30.
[30] A. Garapon, *Justice digitale: révolution graphique et rupture anthropologique*, Paris, 2018, pp. 31-41.
[31] G. D'Acquisto, *Intelligenza artificiale. Elementi,* Torino, 2021, p. 195 ff.
[32] This term was used by E. Frontoni, M. Paolanti, *AI-based decision support system, in* E. Calzolaio (ed.), *La decisione nel prisma dell'intelligenza artificiale,* cit., p. 9.
[33] R. Parasuraman, D. Manzey, *Complacency and Bias in Human Use of Automation: An Attentional Integration*, in *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 52 (3)/2010, p. 390.

among which is the limitation of the human role in the decision-making[34] and thus the risk of losing the humanity that intrinsically characterizes the interpretation of the law[35].

On the other hand, some authors have highlighted that predictive programs have the potential to make the justice system more efficient[36], especially in fields where the rules do not change frequently, and citizens expect uniformity in the application of law[37]. Also, the use of predictive software which foresees the possible outcome of a litigation could push lawyers to reach a deal out-of-court or engage in an alternative dispute resolution (ADR) and thus reduce the workload of judicial systems[38].

Because of the speed and accuracy of their results, these technologies have the power to change the way we view the law (some have described them as *disruptive*[39]), and as lawyers we have to reflect on how to balance the need for uniformity, certainty and speed of trials with the protection of the human rights at stake, such as the right to a fair trial (equality of arms and respect for the adversarial process).

As per usual with new technologies, innovation precedes regulation, and as these predictive tools are already being implemented in many legal systems, the question is not if they should be used, but how to use them in a lawful way which is compatible with the general principles of the rule of law.

In the next paragraphs we will focus on the issues arising from the use of this technology in the judiciary and the solutions provided, both binding and non-binding. We will see that legal systems are regulating predictive justice in different ways and while some of them are more open to its use, others seem to be more cautious. We will argue that educating lawyers and stakeholders about the opportunities and shortcomings of AI is a better solution than banning certain uses of it, also because the utility and effectiveness of these bans is questionable.

2. Shortcomings of algorithmic justice

The use of predictive justice tools can be tricky for many reasons, which can be understood if we take a look at what happened in Wisconsin regarding the use of the COMPAS, a risk assessment tool implemented in many jurisdictions of the United States "to inform decisions regarding the placement, supervision and case management of offenders"[40]. As we can read in the Practitioner's guide: "COMPAS is a statistically based risk assessment developed to assess many of the key risk and need factors in adult correctional populations and to provide information to guide placement decisions". It is used to predict recidivism based on a number of factors, such as substance abuse, family criminality, financial problems, violence history, and others[41].

---

[34] S. F. Ahmad et al., *Impact of artificial intelligence on human loss in decision making, laziness and safety in education*, in *Humanities and Communication,* 10 (1)/2023, p. 4, where the Authors also refer to B. Ghosh et al, *Taking a Systems Approach to Adopting AI (hbr.org)*, in *Harvard Business Review,* May 9, 2019, in stressing that human capabilities such as intuitive analysis, critical thinking and creative problem-solving are getting out of decision-making., which will likely entail their being lost.

[35] L. Vagni, *The Role of Human Judge in Judicial Decisions: preliminary remarks on legal interpretation in the age of Artificial Intelligence*, in E. Calzolaio (ed.), *La decisione nel prisma dell'intelligenza artificiale,* cit., p. 200, where the Author writes: "The scientific and legal knowledge, the capacity to collect, classify and compare data, are important skills for solving the case, but they need to be accompanied by the ability of the judge to interpret law. This is a human ability, as it needs awareness of the contextual dimension of law; in other words, it needs humanity: a free will that impacts with the concrete facts of the case together with the responsibility to seek justice for that case".

[36] R. Simmons, *Big data, machine judges, and the legitimacy of the criminal justice system*, in *University of California Davis Law Review,* 52/2018, pp. 1072-1074.

[37] C. Castelli, D. Piana, *Giustizia predittiva. La qualità della giustizia in due tempi,* in *Questione giustizia,* 4/2018, p. 165.

[38] *Ibid,* p. 163

[39] T. Sourdin, *Judges, Technology and Artificial Intelligence,* Cheltenham, 2021, p. 10 ff.

[40] *Practitioner's Guide to COMPAS Core,* cit., p. 1.

[41] For more info on the COMPAS, see the *Practitioner's Guide to COMPAS Core* issued by Equivant (previously called Northpointe, is the company that owns the tool).

In 2016, an NGO denounced this tool as biased against black people., who, in similar situations, were considered more at risk of recidivism than white people[42]. In the case *State v. Loomis*[43], the Wisconsin Supreme Court held that "a circuit court's consideration of a COMPAS risk assessment at sentencing does not violate a defendant's right to due process ruled in favour of the use of these programs", although it specified that judges had to be cautious while using risk assessment tools[44] as decisions were not to be based solely on their output but supported by other independent factors[45].

This ruling has been criticized by some authors[46], and rightfully so, considering that the Court overlooked the influence that a highly technological tool can have on human decisions and only focused on the accuracy of the data fed to the algorithm, while ignoring its computational procedure, which was not disclosed because of intellectual property rights.

In fact, while the accuracy of the dataset is very important, it is also fundamental to know how much weight each parameter carries in the final output. When it is not possible to access the functioning of the algorithm and the way it processes data (either because of trade secrets or because AI mechanisms are difficult to explain), there is a so-called "black box"[47].

This black box entails a set of negative consequences: first of all, if people do not know how algorithms work, and which parameters are used or their weight in the final decision, they are less willing to trust them[48]. Also, the fact that these algorithms use statistics to determine what should happen to an individual is likely to make people feel uncomfortable[49].

Even when these tools are not protected by trade secrets, the risk of discrimination is not erased. For example, the PATTERN (Prisoner Assessment Tool Targeting Estimated Risk and Needs)[50] is a tool developed by the US Department of Justice; its statistics and functioning are public, but this has not prevented stakeholders from denouncing biases in the tool[51].

It seems, indeed, that discrimination is embedded in the functioning of these programs, considering that they do not have a creative intelligence[52] as their output is based on the processing of existing data. This way of functioning cannot help but repeat past situations - and prejudices - *ad infinitum,* without leaving any room for some kind of evolution.

---

[42] J. Angwin et al, *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks,* May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[43] *State v. Loomis,* 371 Wis. 2d 235.

[44] See par. 8 of the decision: "Ultimately, we conclude that if used properly, observing the limitations and cautions set forth herein, a circuit court's consideration of a COMPAS risk assessment at sentencing does not violate a defendant's right to due process", par. 35 "Although we ultimately conclude that a COMPAS risk assessment can be used at sentencing, we do so by circumscribing its use. Importantly, we address how it can be used and what limitations and cautions a Circuit court must observe in order to avoid potential due process violations". The Court thus held that there was no violation of the due process right because the first instance court did not entirely base its decision on the outcome of the software. Nevertheless, the same Court, in paragraph 47, stated that the right of due process means "right to be sentenced based upon accurate information", which includes "the right to review and verify information contained in the PSI (Presentence Investigation Report) upon which the Circuit court bases its sentencing decision", which was dubiously respected in this case, as neither the judges nor the defendant had access to the algorithm.

[45] See para. 9 of the decision: "We determine that because the Circuit court explained that its consideration of the COMPAS risk scores was supported by other independent factors, its use was not determinative in deciding whether Loomis could be supervised safely and effectively in the community".

[46] A. L. Washington, *How to argue with an algorithm: lessons from the COMPAS-ProPublica debate,* in *Colorado Technology Law Journal,* 17 (1)/2019, pp. 1-37, G. Noto La Diega, *Against the dehumanisation of decision making-Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information,* in *JIPITEC – Journal of Intellectual Property, Information Technology and E-Commerce Law*, 9/2018.

[47] *Ibid*, pp. 9-16.

[48] *Ibid,* p. 5.

[49] R. Simmons, 2018, *Big Data, machine judges, and the legitimacy of the criminal justice system,* cit., p. 1097.

[50] To see more about the PATTERN: BOP: First Step Act, Resources

[51] See the letter written by the Leadership Conference on Civil Rights, https://civilrights.org/resource/comment-letter-to-department-of-justice-on-pattern-first-step-act/

[52] G. D'Acquisto, *Intelligenza artificiale,* cit., p. 195 ff.

AI tools employed for predicting the likelihood of the success of a claim present other types of issues, especially if employed in the decision-making phase.

As stressed above, if used by lawyers, these programs could be a support tool for testing the solidness of their claim; nevertheless, if used by judges, they could be more problematic, especially considering that judges will likely rely a heavily on their outcome.

Also, because of the way they work, the use of these algorithms could lead to a very dangerous "crystallisation"[53] of the law. In fact, judicial decision-making is not a static, simple task; it is a holistic one[54], which needs some skills that are typically human. In each of these tasks, judges will decide not like an automaton, but by interpreting the law and the facts according to their culture and their own personal experience[55] (and this is the reason why sometimes we can see that the interpretation of a statute can vary even if the literal rule remains the same). The humanity of decision-making is fundamental because while legal certainty is important, it is also crucial for the judge to have the ability to "update" the law to the current time[56]. Moreover, a little discretion is fundamental when deciding which facts are relevant for the decision or how to interpret the evidence[57].

One could argue that judges can be influenced by past decisions even without the use of AI; but one thing is to read a decision on a case which is similar to the one at stake, another is having a high performative "intelligent" tool giving percentages of why the claim should be successful or not or the amount of money which should be awarded for damages. At this point, in fact, it would be much more difficult to overlook the outcome of such tool.

As shown above, it would be preferable to let the law "breathe" and adapt to current situations. This does not mean that human interpretation of the law is always correct, but at least it preserves the humanity which intrinsically permeates the decision-making. As has been highlighted: "one of the points about using algorithms is that they should do things better and more safely than humans"[58], and while AI produces excellent results in other fields, this is not the case in legal interpretation.

It must also be considered that whilst human judges can be held accountable for their mistakes[59], the same cannot be said for algorithms[60], and while humans feel the pressure of taking a good decision, because they fear the sanctions[61], this is not the case of AI, which cannot sense as a human being does and certainly does not have any empathy or fear for the adverse consequences of its own decision-making[62].

Besides, if we refer to the Italian system, we can also see that there are some provisions that clarify that judicial decisions should be made case by case by a human judge. One example can be found in Article 116 of the Code of Civil Procedure, which prescribes that judges may infer evidence

---

[53] G. Garapon, *Les enjeux de la justice prédictive, in La semaine juridique,* 1-2/2017, p. 52, where the Author talks about the so-called "*effet moutonnier*", which means that predictive justice leads to conformism in case law.

[54] G. Noto La Diega, *Against the dehumanisation of decision making,* cit., p. 8.

[55] L. Moccia, *Riflessioni sparse (e qualche involontario aforisma) su interpretazione e diritto,* in *Rivista trimestrale di diritto e procedura civile,* 3/2012, p. 918 ff.

[56] L. Rovelli, *Giustizia predittiva. Variazioni sul tema,* in *Contratto e impresa,* 3/2021, p. 745, where the Author talks about the so-called *"innovazioni mature"* (mature innovations) of case law.

[57] A. Carratta, *Decisione robotica e valori del processo*, in *Rivisita di diritto processuale,* 2/2020, p. 504.

[58] J-S. Borghetti, *How can artificial intelligence be defective?,* in R. Schulze, D. Staudenmayer, S. Lohsse (eds.), *Liability for artificial intelligence and the internet of things,* Oxford, 2019, p. 68.

[59] See the Italian law for the civil liability of judges, no. 117 of the 13th of April 1988.

[60] Some Authors have pointed out that behind an algorithm there is always a human decision, thus there is always a human being that could be held accountable for the "decisions" of the algorithm, see G. D' Acquisto, *Intelligenza artificiale,* cit., p. 120, but there is not a widespread consensus on the topic and also the "behaviour" of AI could depend on how it was trained, and the trainer could be different from the engineer who created the algorithm, see U. Ruffolo, *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giurisprudenza italiana,* 7/2019, p. 1692.

[61] G. Noto La Diega, *Against the dehumanisation of decision making,* cit., p. 10.

[62] *Ibid,* p. 11.

from the parties' behaviour. This rule is interesting because it shows that a great deal of value is given to empathy, which is lacking in the technology, at least the one available at the moment[63].

Another example is Article 525, paragraph 2 of the Code of Criminal Procedure, which states that the judicial board making the final the decision has to be the same as the one that participated in the pleading, otherwise the ruling can be considered void. Even this rule gives importance to the personal perception of the judges[64].

Article 220 of the Code of Criminal Procedure also prohibits the use in judicial judgments of criminological reports that state psychological traits (such as the tendency to commit crimes) which do not derive from mental illness[65]. Therefore, if judges cannot delegate a criminological report to another human being, *a fortiori* they should not be able to entrust the profiling of the defendant to a piece of software like a risk assessment tool[66].

Finally, Article 111, paragraph 6 of the Italian Constitution prescribes that every judicial provision has to be motivated by the judge. This rule is important because it shows that, at least in the Italian legal system, software programs could not replace the judge in any part of the decision nor contribute to it to the extent that the judge is no longer able to provide motives for each part. This is the reason why it is more likely that if these tools are to be used in the judicial system, it would be better to implement them for minor tasks and not in the decision-making phase.

Considering all the issues which can arise from the implementation of AI in the judicial system, and taking into account the fact that this technology is available to more and more judges and law firms, regulators have provided for some guidance on how to use them. Hence, it is interesting to investigate how the issues have been tackled in different legal systems using both binding and non-binding frameworks.

3. Ethical principles for the use of AI in judicial systems

In December 2018, the Council of Europe European Commission for the Efficiency of Justice (CEPEJ)[67] published the Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment[68]. When drawing up this document the authors took into account what happened with the COMPAS software tool and were aware of the importance of avoiding discrimination, lack of transparency and a deterministic approach to justice. This was the first document to directly address the issue of the use of AI in judicial systems and is, still nowadays, one of the most important frameworks of ethical principles for AI in the judiciary[69].

The principles are formulated in a broad way, in order to make them easy to adapt to different situations.

---

[63] G. D'Acquisto, *Intelligenza artificiale,* cit*.*, p. 127 ff.

[64] D. Staffieri, at the Conference titled "*Algoritmi, intelligenza artificiale e decisione amministrativa, giudiziaria e tributaria: problemi e prospettive*", held by the Global Professionals for Artificial Intelligence on the 25th of February 2022.

[65] In Italian the article prescribes that *"non sono ammesse perizie per stabilire l'abitualità o la professionalità nel reato, la tendenza a delinquere, il carattere e la personalità dell'imputato e in genere le qualità psichiche indipendenti da cause patologiche".*

[66] D. Staffieri at the aforementioned conference.

[67] See par. 128, 131, 137, 160 of the Ethical Charter.

[68] For the full text https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment, see also C. Barbaro, *Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo? I lavori in corso alla Commissione europea per l'efficacia della giustizia (Cepej) del Consiglio d'Europa,* in *Questione giustizia,* 4/2018, p. 189 ff.

[69] For comparative studies on ethical frameworks, see, L. Hedler, *Risk and danger in the introduction of algorithms to courts: A comparative framework between EU and Brazil*, in *Oñati Socio-Legal Series*, 2024; G. Lupo, *The ethics of Artificial Intelligence: An analysis of ethical frameworks disciplining AI in justice and other contexts of application*, in *Oñati Socio-Legal Series*, 12 (3)/2022, pp. 614-653.

The first principle stresses the importance of the respect for the fundamental rights listed in the European Convention on Human Rights (ECHR). This means that AI tools "should also be used with due respect for the principles of the rule of law and judges' independence in their decision-making process" and should "not undermine the guarantees of the right of access to the judge and the right to a fair trial".

The second principle focuses on the need to prevent discrimination between individuals or groups of individuals. This means that the stakeholders involved must ensure that the use of predictive justice does not create or aggravate discrimination, especially when the processing is based on "sensitive data", as in the COMPAS case.

The third principle addresses the need for quality and security with regard to the processing of judicial decisions and data, which has to be done using certified sources. This principle is particularly important because AI is "fed" with a huge amount of data, which are elaborated by the algorithm and then processed in order to give an outcome[70]. However, this outcome is profoundly influenced by the quality of the data; with the consequence that if the said data are somehow biased, the output of the algorithm will also be biased[71].

The fourth principle highlighted by the CEPEJ is that data processing has to be accessible and transparent in order to reach a balance between the protection of intellectual property and the need for transparency.

The last principle precludes a prescriptive approach and ensures that users are informed actors and in control of their choices (i.e. professionals who use these AI tools should, at any moment, be able to review judicial decisions and the data employed to produce a result).

In order to make it easier for stakeholders to implement these principles, the CEPEJ's Artificial Intelligence Advisory Board is soon going to be publishing a tool to provide for clear and practical guidance on how to apply the Ethical Charter in practice[72].

The CEPEJ principles were followed in 2019 by the Ethics Guidelines for Trustworthy AI[73], published by the European Commission, and written by the High-Level Expert Group on AI.

In the guidelines, the experts address the need for AI to be human-centric[74] and the fact that, in order to be trustworthy, AI has to be lawful, ethical and robust both from a technical and social perspective.

In particular, AI has to respect the fundamental rights as enshrined in the EU Treaties[75] and has to comply with four ethical principles, which are the respect for human autonomy, the prevention of harm, fairness and explicability.

The guidelines also state that AI, to be trustworthy, has to comply with some requirements, such as accountability, non-discrimination, technical robustness, privacy and data governance, transparency and human agency and oversight, all while guaranteeing human wellbeing.

In 2021, building on the 2019 framework, the guidance Ethics by Design and Ethics of Use Approaches for Artificial intelligence was published[76]. This work, drafted by a panel of experts and requested by the European Commission, offers guidelines for adopting an ethical approach while designing, developing and deploying and/or using AI based solutions. This document, like the one from 2019, aims to offer assistance when adopting an ethics by design approach.

---

[70] G. D'Acquisto, *Intelligenza artificiale,* cit., p. 127 ff.

[71] A. Daly, S. K. Devitt, M. Mann, *AI Ethics Needs Good Data*, *in* P. Verdegem (ed.), *AI for Everyone?: Critical Perspectives,* London, 2021, pp. 103-121, where the Authors stress the importance of "good data" for the training of AI algorithms.

[72] For more information on this tool, see https://www.coe.int/en/web/cepej/-/cepej-artificial-intelligence-advisory-board-aiab-.

[73] For the full text see https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html.

[74] High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*, p.4.

[75] *Ibid,* p. 11.

[76] https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf.

Although the aforementioned sets of principles have been a step forward in recognizing the opportunities and adverse consequences of using AI in judicial systems, they cannot be the only means of addressing the challenges posed by AI.

First of all, the broad definitions expressed in the principles have to be implemented in practical tools, and this is where the first issue arises[77]. In fact, while frameworks excel in identifying ethical issues, they are less convincing in providing practical recommendations. From a technical point of view, it is difficult to implement these broad ethical principles which are neither concrete nor technical[78], and only in some fields are engineers and other academics conducting interesting research to enshrine broad principles in practical tools, i.e. the field of "privacy preserving" is nowadays a subfield of machine learning, where researchers focus on developing programs that respect the privacy of the stakeholders involved[79].

It should also be noted that while these ethical frameworks focus a great deal on broad principles, they are not able to address what happens when there is a breach and AI is used for illicit purposes since they are non-binding, thus cannot impose sanctions.

Therefore, we share the opinion that "principlism" alone is not able to achieve full ethicality or in general to prevent and assess any negative consequence arising from the deployment of such tools, if it is not put into practice using other means[80].

4. A "hard law" approach to predictive justice: the examples of the European Union, France, and Italy

As can been seen from the frenetic legislating activity of the last few years, the European Union is trying to regulate almost every interaction between law and technology through new legislation or by modifying pieces of legislation already in force. Some authors have even pointed out that the continuous introduction of new statutes and reforms makes it difficult for interpreters, scholars or citizens in general to find and apply the right rules[81].

More specifically, as regards the use of AI software in the judiciary, we can see that the General Data Protection Regulation (GDPR) [82] and the so-called AI Act[83] establish some interesting rules.

Article 22 of the GDPR prescribes that "the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling[84] which produces legal effects

[77] E. Prem, *From ethical AI frameworks to tools: a review of approaches,* in *AI and Ethics*, 3/2023, pp. 699–700.
[78] *Ibid,* p. 702.
[79] *Ibid,* p. 710.
[80] *Ibid*, pp. 711–712. Moreover, other problems arise from the fact that research has also highlighted how the most influential ethical guidelines come from the more economically developed areas of the world, with a strong underrepresentation of other areas, such as Africa, meaning that the more economically developed countries are shaping the terms of the debate more than others, which raises concerns about cultural pluralism and the demands of global fairness. Also, sometimes we can see a contrast in the application of the broad principles stated in those ethical frameworks. For example, the need for a larger dataset for the accuracy of the output can contrast with the respect for the privacy and autonomy of the data subject, see A. Jobin, M. Ienca, E. Vayena, *The global landscape of AI ethics guidelines,* in *Nature Machine Intelligence,* 1/2019, pp. 389–399.
[81] E. Bellisario, *Il pacchetto europeo sulla responsabilità per danni da prodotti e da intelligenza artificiale. Prime riflessioni sulle Proposte della Commissione,* in *Danno e resp*., 2/2023, p. 154.
[82] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
[83] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts, Brussels, 21.4.2021, as approved by the European Parliament on the 6th of March 2024 and then by the European Council on the 21st of May 2024.
[84] According to art. 4 GDPR "'profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements".

concerning him or her or similarly significantly affects him or her", from which it can be inferred that purely automated judicial decisions are forbidden.

The article makes a few exceptions to this general rule, namely when the automated decision making is necessary for entering into, or performance of, a contract between the data subject and a data controller; when the automated decision making is authorized by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests or is based on the data subject's explicit consent. In the cases where the first and third exception apply, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision. In cases where solely algorithmic automated decisions are allowed, they "shall not be based on special categories of personal data referred to in Article 9(1)" (sensitive data on health, political or religious belief, race, etc..), unless there is the explicit consent of the data subject (Art. 9 (2) point a) or the profiling is necessary for reasons of substantial public interest and "suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place". Whether risk assessment tools such as the COMPAS could be included in the hypothesis of substantial public interest is debatable.

In case of breach of Art. 22, the supervisory authority shall ensure the imposition of administrative fines[85] on the controller[86] or the processor of the data[87].

Article 22 of the GDPR is an important rule and it was the first example of the explicit distrust of the European legislator towards fully automated decision making. However, its wording is quite unclear, especially when it refers to decisions based *solely* on automated decision-making. For example, if the algorithmic system makes the decision and then a human being reviews it, would the decision still be purely automated?[88] What is the necessary human input for the decision to be considered still human? And even if a human being reviews the decision, how is it possible to make sure that the final decision is actually taken by the judge and has not been overshadowed by the outcome of the data processing?

Lately, the European Union has taken a step further, specifically for the regulation of the use of risk assessment tools. As stated in Article 5, subparagraph d) of the final version of the AI Act of the 6th of March 2024, which was approved by the European Parliament on the 13th of the same month and then by the European Council on the 21st of May 2024, among the prohibited AI practices there is "the placing on the market, the putting into service for this specific purpose, or the use of an AI system for making risk assessments of natural persons in order to assess or predict the likelihood of a natural person committing a criminal offence, based *solely* on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity". Non-compliance with the prohibition of the AI practices referred to in Article 5 entails administrative fines of up to 35 000 000 EUR or, if the offender is an undertaking, up to 7 % of its total worldwide annual turnover for the preceding financial year, whichever is higher[89].

Therefore, similarly to the wording of the GDPR, the AI Act prescribes that the risk of natural persons committing a crime cannot be assessed or predicted *based solely* on their profiling or

[85] See Art. 83 GDPR.
[86] According to Art. 4 GDPR, "'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".
[87] According to Art. 4 GDPR, "'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller".
[88] G. Noto La Diega, *Against the dehumanisation of decision making*, *cit.,* p. 19.
[89] Art. 99, par. 3 AI Act.

assessing their personality traits and characteristics. This should mean, *a contrario,* that these tools could still be used if combined with other factors in a human decision.

The rule is at best useless, as the ban on *solely* automated decision-making (concerning not only criminal risk assessment, but any automated processing activity) could already be inferred by Art. 22 GDPR. Anyway, the problem with the use of AI tools in delicate activities is that, as has been stressed above, humans tend to rely on the outcome of highly sophisticated computer software, thus it would be difficult for a human judge to depart from the outcome of a highly technical tool, especially when it shows through numbers and statistics why a certain decision should be made.

In France, a country which has traditionally been sensitive to the interactions between law and new technologies[90], Article L111-13 of the *Code de l'organisation judiciaire* was introduced by the *Loi* n° 2016-1321 of the 7th October, ensuring free access to all judicial decisions given by all jurisdictions in electronic format[91]. Moreover, to protect the privacy of the parties involved, the decisions are anonymized[92].

The free access to judicial decisions has been a turning point for the development of predictive justice algorithms, as the number of available decisions can be used as big data to train the algorithms.

Furthermore, to protect judges and the independency of their work, in 2019[93] another paragraph was added to Article 111-13, specifying that it is prohibited to profile judges in order to evaluate, analyse, compare or predict their professional practices. The violation of this provision is a criminal offence punishable with up to five years in prison[94]. This statute, as the first example of such a ban in the world[95] is clearly aims to prevent anyone, especially the so-called LegalTech companies, from revealing the pattern of judges' behaviour in court decisions[96], although it has been criticized by some authors, as they have held that this provision could be read as being in violation of the freedom of expression in Article 10 of the ECHR[97].

In addition to these rules, France even introduced a stricter regime for automated decision-making based on the profiling of a person. In fact, in 2018[98], Art. 10 of the 1978 *Loi informatique et libertés* was modified to add a rule establishing that no judiciary decision implying the evaluation of the behaviour of a person can be based on the automated processing of personal data to decide on some aspects of their personality. Nowadays, following a reform[99], the same principle has been

---

[90] G. Resta, *Governare l'innovazione tecnologica. Decisioni algoritmiche, diritti digitali e principio di uguaglianza,* in *Pol. Dir.,* 2/2019, p. 228.

[91] The current text, as modified in 2019, prescribes : *"Sous réserve des dispositions particulières qui régissent l'accès aux décisions de justice et leur publicité, les décisions rendues par les juridictions judiciaires sont mises à la disposition du public à titre gratuit sous forme électronique.*
Les nom et prénoms des personnes physiques mentionnées dans la décision, lorsqu'elles sont parties ou tiers, sont occultés préalablement à la mise à la disposition du public. Lorsque sa divulgation est de nature à porter atteinte à la sécurité ou au respect de la vie privée de ces personnes ou de leur entourage, est également occulté tout élément permettant d'identifier les parties, les tiers, les magistrats et les membres du greffe.
Les données d'identité des magistrats et des membres du greffe ne peuvent faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées. La violation de cette interdiction est punie des peines prévues aux articles 226-18,226-24 et 226-31 du code pénal, sans préjudice des mesures et sanctions prévues par la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.
Les articles L. 321-1 à L. 326-1 du code des relations entre le public et l'administration sont également applicables à la réutilisation des informations publiques figurant dans ces décisions.
Un décret en Conseil d'Etat fixe, pour les décisions de premier ressort, d'appel ou de cassation, les conditions d'application du présent article" .

[92] *Ibid.*

[93] The reform was introduced by *Loi* n° 2019-222.

[94] See Articles 226-18, 226-24 and 226-31 of the French Criminal Code.

[95] A. Duflot, *Artificial intelligence in the French Law of 2024,* in *Legal Issues in the Digital Age,* 5 (1)/2024, p. 44.

[96] *Ibid.*

[97] *Ibid.*

[98] *Loi* n. 2018-493.

[99] *Ordonnance* n°2018-1125.

transposed to Art. 47 of the same law[100]. Contrary to the wording of the GDPR and the AI Act, here it seems that AI could not be used at all, otherwise the provision would have specified that it only referred to decisions based *solely* on automated processing of personal data. This means that the use of predictive tools similar to the COMPAS case would not be possible in France[101], although this rule does not prevent the use of predictive algorithms in other fields outside the profiling, for example for the calculation of damages[102].

In the light of the foregoing, it seems that France has opted for a stricter regime for risk assessment tools than that laid down by the European Union.

Lately, even Italy has taken the first steps towards creating a national legal framework for AI. In fact, on the 23rd of April 2024, the Italian Government voted on the text of a proposal for an upcoming law on AI[103]. Even if the legislative process is only at its early stages (it still has to be scheduled to be discussed and approved by both of the Chambers of the Italian Parliament[104]), it is interesting to analyse some of its provisions since they are relevant for the regulation of the use of predictive justice tools.

First of all, Art. 12 of the proposal prescribes that the use of AI by intellectual professionals is permitted only for support activities, while the intellectual part of the work has to be prevalent. Moreover, the professional must inform the client that s/he uses AI. This article can be applied to the work of lawyers who want to employ predictive justice tools in their routine.

Even more interesting for the purposes of this work is Art. 14, which prescribes that AI tools can be used in the judiciary only for organizational purposes, to simplify judiciary work or for research on case law or doctrine. The Ministry of Justice is going to clarify how to deploy AI systems in the judicial system.

---

100 *"Aucune décision de justice impliquant une appréciation sur le comportement d'une personne ne peut avoir pour fondement un traitement automatisé de données à caractère personnel destiné à évaluer certains aspects de la personnalité de cette personne.*
Aucune décision produisant des effets juridiques à l'égard d'une personne ou l'affectant de manière significative ne peut être prise sur le seul fondement d'un traitement automatisé de données à caractère personnel, y compris le profilage, à l'exception : 1° Des cas mentionnés aux a et c du 2 de l'article 22 du règlement (UE) 2016/679 du 27 avril 2016, sous les réserves mentionnées au 3 du même article 22 et à condition que les règles définissant le traitement ainsi que les principales caractéristiques de sa mise en œuvre soient communiquées, à l'exception des secrets protégés par la loi, par le responsable de traitement à l'intéressé s'il en fait la demande; 2° Des décisions administratives individuelles prises dans le respect de l'article L. 311-3-1 et du chapitre Ier du titre Ier du livre IV du code des relations entre le public et l'administration, à condition que le traitement ne porte pas sur des données mentionnées au I de l'article 6 de la présente loi. Ces décisions comportent, à peine de nullité, la mention explicite prévue à l'article L. 311-3-1 du code des relations entre le public et l'administration. Pour ces décisions, le responsable de traitement s'assure de la maîtrise du traitement algorithmique et de ses évolutions afin de pouvoir expliquer, en détail et sous une forme intelligible, à la personne concernée la manière dont le traitement a été mis en œuvre à son égard.
Par dérogation au 2° du présent article, aucune décision par laquelle l'administration se prononce sur un recours administratif mentionné au titre Ier du livre IV du code des relations entre le public et l'administration ne peut être prise sur le seul fondement d'un traitement automatisé de données à caractère personnel".

101 G. Resta, *Governare l'innovazione tecnologica,* cit., p. 228.

102 Commission nationale consultative des droits de l'homme, *Avis relatif à l'impact de l'intelligence artificielle sur les droits fondamentaux,* in *Journal officiel de la république française n° 0091 du 17/04/2022,* par. 28, where it is stated that "*d'ores et déjà ce type de recours à l'IA est interdit en France puisqu'aucune décision de justice impliquant une appréciation sur le comportement d'une personne ne peut avoir pour fondement un traitement automatisé de données à caractère personnel destiné à évaluer certains aspects de la personnalité de cette personne ". La formule ainsi retenue n'écarte cependant pas toute possibilité de fournir aux magistrats une application d'IA répondant à d'autres finalités, par exemple pour automatiser le calcul de l'indemnisation d'un préjudice*".

103 In Italy, according to Art. 71 of the Constitution, the Government can propose bills that then have to be approved by both of the legislative Chambers.

104 See Arts. 70 and 72 of the Italian Constitution. At the moment, the proposal has been authorized by the President of the Republic to be presented to one of the Chambers on the 17th May 2024, (see the web page of the President of the Republic: https://www.quirinale.it/elementi/112446), and it was presented to the Senate on the 20th May 2024, Atto del Senato no. 1146, see the web page of the Senate: https://www.senato.it/leg/19/BGT/Schede/Ddliter/testi/58262_testi.htm.

Paragraph 2 of the article specifies that it is always up to the human judge to decide on the interpretation of the law, the evaluation of facts or proof and on every court order.

If adopted, this provision would preclude the use of e-discovery software such as NUIX, which was used by the Public Prosecutor of Genoa to examine a huge number of digital documents in the investigations following the collapse of the Morandi Bridge in 2018[105]. NUIX is not strictly a predictive justice tool, according to the definition we gave above, but it is interesting as an example of the various ways in which AI can be used in trials. It combines machine learning and natural language processing techniques to identify, among many documents, the elements that are relevant for the investigation[106]. Following the use of this piece of software by the Prosecutor, the defence attorneys held that there had been a violation of the equality of arms, as they did not have a similar instrument to analyse the documents in question. The judges, however, rejected the motion[107], on the ground that the defence did not ask to use the tool, adding that they could also download a similar one from the internet. Assuming that this could be accepted as a correct response to a serious problem affecting the right to a fair trial, the judges failed to consider that NUIX software, just like COMPAS, is a proprietary software (it belongs to a private Australian company). Therefore, the defence could not access the algorithm and consequently it was not known how it processed data and which elements were prioritised in this serious case.

The new proposal, if adopted, will likely ban this use of AI software in the judiciary, unless the prevalent interpretation of the rule is that the human factor in the decision still persists even when judges utilise an AI tool to take their decision, but this interpretation seems to depart too much from the wording of the provision (if retained in the final text).

The above-mentioned proposal does not include a sanction for the infraction of this ban, although it states that the Italian Government should be delegated to adopt one or more pieces of legislation to define the sanctions in case of illicit use of AI.

In general, not very unlike the French rules on the topic, the Italian approach shows a very specific distrust of the use of predictive justice software in the decision-making.

These approaches give rise to some issues in relation to EU law. In fact, the rule set out in Art. 5 of the AI Act seems to ban the use of risk assessment tools only when their use is not supported by a human decision-maker, while France and Italy seem to be oriented towards the banning of every use of AI in the decision-making phase, whether or not it is merely used as a tool in the hands of human judges.

Nevertheless, we must question if this approach is allowed by the European regulation, seeing that the very first sentence of the first article states that the purpose of the Act is "to improve the functioning of the internal market". This seems to be alluding to the fact that the AI Act is a maximum harmonisation tool. In fact, if it were not considered as such, and Member States could introduce harsher bans than the ones prescribed by the Regulation, it would be very difficult for a producer of an AI risk assessment tool to be able to sell its products in some States (such as France or Italy) while finding it easy to trade them elsewhere within the EU, with serious impacts on the functioning of the internal market.

In addition to this, one could question whether the decision of banning certain uses of AI in the judiciary by introducing a conspicuous number of new pieces of legislation is really helpful.

First of all, the proliferation of pieces of legislation on the matter generates a great deal of confusion, especially when the use of these new tools is transnational, as technology knows no territorial border. The legal expert who has to study this subject is bewildered by the many legal sources coming not only from the European Union, but also from Member States, all of which leads

[105] Software programs embedding natural language processing technologies for e-discovery are used in many jurisdictions, for example in the UK, their use has been approved by the decision *Pyrrho investments ltd v mwb property ltd* [2016] EWHC 256 (Ch).
[106] More information about the functioning of this tool at https://www.nuix.com/technology/nuix-discover.
[107] Ordinanza del Tribunale di Genova, prima sezione penale, secondo collegio, RGNR. 10468/2018.

to a very intricate legal framework. At some point, some technological company could even decide to take their business to less complicated and more populated markets.

This does not mean that every technological practice should be permitted, but at this stage we have so many AI rules in the European territory and so few incentives for tech companies (if compared to other countries, such as the United States) that there is the ever-present risk of hindering innovation.

Moreover, it must be noted that, as the spread of new technologies does not recognize physical borders and information can travel fast through social media, bans risk being at best useless (see what happened when the Italian Data Protection Authority tried to ban the use of ChatGPT: soon after people started using other generative tools with similar performances or devices enabling them to localise their position elsewhere outside of Italy in order to have access to ChatGPT).

In addition to this, for predictive justice tools specifically, it might also be difficult to recognize just from a final decision if the judge has delegated some parts, or the totality of it, to an AI tool, with serious consequences in terms of application of the afore-mentioned rules.

At this point, we must question whether binding rules are the best way to deal with the shortcomings of the use of predictive justice tools, as it could be argued that the problem is one of legal education: if lawyers, judges and any other legal professionals understand not only the importance of humanity in their everyday work but most importantly the basics of the functioning of AI, they will be more aware of the tasks which should be delegated to it and those which should, preferably, be carried out by humans.

5. Guidelines, education, and corporations: the approach of the United States

In the search for a balance between innovation and regulation, the United States, unlike Europe, is avoiding introducing hard law measures to cope with the challenges provided by AI.

One reason could be the fact that the United States is a great hub for AI companies and LegalTechs (which are mainly situated in Silicon Valley, California), and seems to be eager to maintain its central global position for AI innovation in respect to other international competitors in the AI field (i.e. China, which is the second largest leader in the so-called AI race)[108]. To do so, it is steering clear of hard law regulations so that AI companies can innovate the field without having to deal with burdensome regulations.

In order to maintain its predominance over AI, the United States is also pursuing the strategy of involving corporations in the regulation of AI; in fact, several representatives of the most important AI tool companies have been invited to Congress to discuss the opportunity of future AI regulations. One significant meeting was the one held on the 13th of September 2023, where the different parties discussed the dangers and challenges of this new technology[109].

It must also be noted that, in the absence of a binding regulation, several corporations have already published their own ethical guidelines on how to implement AI[110]. However, one must point out that delegating the regulation of AI to industry stakeholders is not the best option for several reasons. First of all, their frameworks have no democratic legitimacy[111]; secondly, there is the risk of creating a regulatory environment made up of different perspectives which lack consistency[112];

---

[108] Y. Walter, *Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation – A contemporary overview and an analysis of socioeconomic consequences,* in *Discover Artificial Intelligence,* 4/2024, p. 7.

[109] https://edition.cnn.com/2023/09/13/tech/schumer-tech-companies-ai-regulations/index.html.

[110] See Meta Guidelines on Responsible AI, https://ai.meta.com/responsible-ai/, where they call for the implementation of the principles of privacy and security, fairness and inclusion, robustness and safety, transparency and control, accountability and governance. See also Open AI Safety standards: https://openai.com/safety-standards.

[111] Y. Walter, *Managing the race to the moon,* cit., pp. 11-12.

[112] *Ibid.*

finally, and most importantly, corporations do not operate for societal development but to pursue their own profit[113].

This means that a proper regulation should come from democratic organs. In this sense, the President of the United States has taken an important step through the Executive Order[114] on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence of the 30th of October 2023[115], where among the many topics tackled, there is also the use of AI in the judicial system. Here, the challenges of AI are addressed through the request to establish non-binding guidelines which can help face the issues deriving from the implementation of AI in proceedings, without hindering innovation through hard law.

More specifically, according to section 7 of the executive order, and to address unlawful discrimination and other harms that may be exacerbated by artificial intelligence, the Attorney General[116] shall coordinate with and support agencies in their implementation and enforcement of existing Federal laws to address civil rights and liberties violations and discrimination perpetrated through the use of artificial intelligence.

The Attorney General was also urged to direct the Assistant Attorney General in charge of the Civil rights division to convene a meeting, within 90 days from the date of the order, of the heads of the Federal civil rights offices to discuss how to prevent and address algorithmic discrimination in the judiciary, increase coordination between agencies and improve the external stakeholder engagement to promote public awareness of potential discriminatory uses and effects of AI tools.

This meeting took place on the 11th of January 2024[117]. During the session, the attendees discussed how to protect civil rights, not only through policy initiatives but also education and outreach. In particular, participants emphasized the importance of educating the public about the way in which AI is able to violate federal protections and the need to develop holistic remedies to address these harms.

According to the executive order, moreover, the Attorney General is invited to consider providing guidance, technical assistance and training to State, local, Tribal, and territorial investigators and prosecutors on best practices for investigating and prosecuting civil rights violations and discrimination related to AI.

To pursue the equitable treatment of individuals, the Attorney General shall, in consultation with the Secretary of Homeland Security and the Director of the Office of Science and Technology Policy submit a report to the President, within 365 days of the date of the order, addressing any use of AI in the criminal justice system, including: sentencing, parole, supervised release, probation, pretrial release or detention and risk assessment.

Also, in order to advance the presence of appropriate technical experts and expertise among law enforcement professionals, within 270 days of the order, the Attorney General shall also consider the best practices, and if necessary, develop recommendations for States, local, Tribal, and territorial law enforcement agencies and criminal justice agencies seeking to recruit, hire, train, promote

---

[113] *Ibid.*

[114] It has been common practice for U.S. Presidents to issue executive orders. The Government of the United States functions on the basis of the separation of powers, which is maintained through a system of checks and balances in the Constitution. The executive power is "vested in a President of the United States of America" by Article II of the Constitution. This article names the President as "Commander in Chief of the Army and Navy" and delegates to him the authority to "take Care that the laws be faithfully executed." As already stated, "nowhere in the Constitution is there any specific reference concerning the power of the President to issue executive orders. Irrespective of this lack of constitutional sanction, all residents since Washington have issued orders and directives which could be technically classified as executive orders", see W. Hebe, *Executive Orders and the Development of Presidential Power*, in *Villanova Law Review*, 17 (4)/1972, p. 688.

[115] Executive Order 14110.

[116] "The Attorney General represents the United States in legal matters generally and gives advice and opinions to the President and to the heads of the executive departments of the Government when so requested. In matters of exceptional gravity or importance the Attorney General appears in person before the Supreme Court", see: https://www.justice.gov/ag.

[117] https://www.justice.gov/opa/pr/readout-justice-departments-interagency-convening-advancing-equity-artificial-intelligence.

and retain highly qualified and service-oriented officers and staff with relevant technical knowledge (such as machine learning engineers, data scientists, data privacy experts, and other highly qualified professionals).

This final provision is particularly important because it incentivizes cooperation and multidisciplinarity, which is probably a better way to deal with the challenges of new technologies than banning their use through a chaotic framework of binding rules, whose effectiveness is to be questioned[118].

8. Final remarks

The use of predictive justice tools has spread in law firms and courtrooms all over the world, causing many problems in terms of the risk of discrimination, lack of transparency and crystallisation of the law. Therefore, it is interesting to reflect on the benefits, shortcomings and the risk-balance of its deployment in such environments.

On the one hand, AI can be used to facilitate some repetitive tasks which do not require a human interpretation of the law or the relevant facts[119], for example to summon the parties or to automatically produce some documents[120] (i.e. when the hearing is being delayed), or to verify if a procedural condition has been met (in the Italian system, in some subjects the completion of an ADR procedure is a condition for filing a claim before a court; AI programs could be used to verify if this condition has been fulfilled).

On the other hand, it would be better to avoid using these tools for decision-making purposes, such as the interpretation and application of the law or the calculation of the rate of recidivism, since the outcome of AI tools can be biased. This can happen because the algorithm considers data in a certain way, or the data used are not representative of the reality, or they are, but in a way that recreates past discrimination[121] (let us go back to the COMPAS example, here the algorithm *per se* was not designed to discriminate against people, nevertheless it took into consideration some aspects of people's lives that were likely to perpetrate past discrimination between black and non-black defendants[122]). Given how AI algorithms work, much attention has to be paid to the way they are designed and the data they are fed.

Moreover, even if they were not biased, AI tools can be problematic because they carry a serious risk of crystallisation of the law or past situations. In fact, as they process and elaborate their output according to the data fed in (which are made up of past judicial decisions), there is the risk of impeding the evolution of law, and this would be inadvisable. In fact, law has to be able to breathe and evolve according to the changing times. As Holmes stated: "we do not realize how large a part of our law is open to reconsideration upon a slight change in the habit of the public mind"[123].

---

[118] In England and Wales, the *Guidance for responsible use of AI in Courts and Tribunals*, issued by the Lady Chief Justice of England & Wales, the Master of the Rolls, the Senior President of Tribunals and the Deputy Head of Civil Justice on 12 December 2023 states that before using AI, judges should have a basic understanding of it. In particular, judges should know that public chatbots do not provide answers from authoritative databases and are mainly trained on U.S. materials. In addition, judges are warned not to enter confidential information in such databases. Nevertheless, the guidance recognizes that judges are not obliged to describe their research or preparatory works to produce a judgment, therefore they are free to use AI. However, among the tasks that should not be performed with the aid of AI there are legal analysis or reasoning.

[119] After all, as stated by Justice Roberts, Chief Justice of the United States, in his *2023 Year-End Report on the Federal Judiciary*, AI will not make judges obsolete, but it will likely change the way they work (although legal profession is notoriously adverse to changes). Justice Roberts himself recognized that legal determinations often involve grey areas that require human judgment that cannot be fully replaced by machines.

[120] S. Abiteboul, F. G'Sell, *Les algorithmes pourraient-ils remplacer les juges?,* in *Le Big Data et le droit*, *Thèmes et Commentaire*, 2019, p. 14.

[121] S. K. Katyal, *Private Accountability in the Age of Artificial Intelligence,* in *UCLA Law Review,* 54/2019, pp. 89-91. See also L. Vagni, *The Role of Human Judge in Judicial Decisions,* cit., p. 189.

[122] G. Resta, *Governare l'innovazione tecnologica,* cit., p. 216.

[123] O. W. Holmes, *The path of the Law*, cit., p. 466.

Another issue that arises with the use of AI tools in the decision-making phase is the risk of relying too much on their outcome: in fact, research shows that users might tend to overestimate the performance of automated aids[124], thus it is particularly difficult for a human being to depart from the conclusions of a highly automated tool. This could lead to humanity getting out of decision-making[125] which is undesirable, since, as has been stressed above, AI tools are not able to interpret law and apply it in the same way as a human judge[126].

Given the many questions that arise from the use of predictive tools in the judiciary, the CEPEJ and the European Commission have admirably tried to outline a set of ethical principles to be followed when using AI software in judicial systems, but they are voluntary generic frameworks, whose effectiveness is limited.

The European Union, first with the GDPR and then with the AI Act, recognized the need to address the challenges posed by these tools, but its answers seem weak because the statutes do not clarify the specific phases in which predictive tools could be used and those in which they should not. France has adopted national rules, and they seem to be pursuing an approach that is even stricter than the EU (the same path is being followed by Italy through a recent legislative proposal). Nevertheless, we are not sure that derogating the European Regulation on this matter is possible, given the way in which AI is formulated to achieve maximum harmonization.

Moreover, one should question whether the European approach of so thoroughly regulating the use of AI is the best answer to the challenges it poses, given how easy it is for people to use technologies even if they have been banned.

At this point one could argue that while it is obvious that these tools raise many concerns, maybe a better answer is to invest in the legal education of the stakeholders and the general public, making them aware of the functioning of AI and the side effects that it can have on the development of law, or the lack of it, and the perpetration of past discriminations.

This is the approach pursued by the United States. Here, the strategy set out in the Executive order of 2023 is to implement "soft law" instruments such as guidelines, reports and best practices.

Interestingly, the executive order advances the proposal for the Advocate General to set out a framework of best practices, and if necessary, develop recommendations to advance the presence of relevant technical experts and expertise on AI among law enforcement professionals. The cooperation between different expertise is in fact imperative when such technical tools are implemented in the legal field, so that lawyers understand the opportunities and shortcomings deriving from the deployment of AI.

Although it is difficult to compare the two different ways of regulating AI, since they are both in their infancy, it is interesting to note that elsewhere in the world the approach is not that of legislating any deployment of AI but making the stakeholders aware of the functioning and possibilities of the technology, so that they can make an informed choice on how (and whether) to implement it.

Abstract

The availability of artificial intelligence software is increasingly permeating the law. In particular, predictive justice has become a popular topic among legal scholars in the last few years because of the availability of artificial intelligence programs allowing users to predict the outcome of a trial or the risk of recidivism during the trial or pre- and post-trial phases. These tools raise many issues in terms of discrimination, lack of transparency and crystallisation of the law, especially because studies show how difficult it is for a "human"

[124] R. Parasuraman, D. H. Manzey, *Complacency and Bias in Human Use of Automation,* cit., p. 392.
[125] B. Ghosh et al, *Taking a Systems Approach to Adopting AI,* cit.
[126] S. Abiteboul, F. G'Sell, *Les algorithmes pourraient-ils remplacer les juges?,* cit., p. 12 ff.

judge to depart from the outcome of a highly technological tool. This paper reviews some examples of the implementation of predictive justice programs in different legal systems, as well as the regulatory instruments that have been introduced in order to reduce the risks generated by their use. To this end, documents containing ethical principles have been adopted, although their effectiveness is limited by their non-binding and generic nature, which make them difficult to embed in artificial intelligence tools. Therefore, some strict statutes have been introduced by the European Union and some Member States. These statutes seem to prohibit some uses of predictive justice tools, although it is doubtful that this is the best solution, since these bans can be easily circumvented. Therefore, we argue that the preferable solution is to stimulate cooperation between technical experts and legal experts, to promote greater awareness among them of the applications and implications of artificial intelligence in judicial proceedings, a solution that is already being pursued in the United States.

Key words: predictive justice, artificial intelligence, risk assessment tools, European Union, United States

*

La disponibilità di software di intelligenza artificiale permea sempre di più il mondo del diritto. In particolare, il tema della giustizia predittiva ha conosciuto grande popolarità tra i giuristi negli ultimi anni grazie alla disponibilità di software che consentono all'utilizzatore di predire l'esito di un procedimento o di calcolare il rischio di recidiva durante il processo o nelle sue fasi antecedenti o posteriori. Questi strumenti pongono serie problematiche in termini di discriminazione, mancanza di trasparenza e cristallizzazione del diritto, atteso che ci sono studi che dimostrano la difficoltà del giudice "umano" nel discostarsi dalle risultanze prodotte da uno strumento altamente tecnologico. Nel contributo si passano in rassegna alcuni esempi di implementazione di programmi di giustizia predittiva in diversi ordinamenti giuridici, nonché gli strumenti di regolamentazione che sono stati introdotti al fine di ridurre i rischi generati dal loro utilizzo. A tal fine, sono stati innanzitutto adottati documenti contenenti principi etici, che però incontrano dei limiti intrinseci dati dalla loro natura non vincolante e dalla loro portata generale, che si esplicano in una certa difficoltà nel tradurli in concreto negli strumenti di intelligenza artificiale. Nell'Unione europea e in alcuni Stati membri sono state adottate delle soluzioni legislative contenenti regole particolarmente stringenti, che sembrano vietare alcuni utilizzi degli strumenti di giustizia predittiva, benché sia lecito dubitare dell'effettività di tale scelta, anche in virtù del fatto che tali divieti possono facilmente essere aggirati. Pertanto, si ritiene che la soluzione preferibile sia quella di stimolare la cooperazione tra esperti tecnici e giuristi, incoraggiando una maggiore consapevolezza di questi sulle applicazioni e le relative implicazioni dell'intelligenza artificiale nel procedimento giudiziario, soluzione che viene già perseguita negli Stati Uniti.

Parole chiave: giustizia predittiva, intelligenza artificiale, strumenti di valutazione del rischio, Unione europea, Stati Uniti