THE IMPACT OF ARTIFICIAL INTELLIGENCE IN THE CONCLUSION OF THE PLEA AGREEMENT

C.S. PARASCHIV

Carmen Silvia PARASCHIV
Professor PhD. "Titu Maiorescu" University of Bucharest, Faculty of Law
E-mail: paraschivcrmn@yahoo.com
ORCID ID: https://orcid.org/0009-0002-1235-5813

Abstract

The digital transformation of the judicial system has made artificial intelligence (AI) an increasingly important vector in the analysis of evidence, in the formulation of plea agreements, and in the anticipation of judicial outcomes (Susskind, 2019; Surden, 2020). The article investigates the impact of these technologies in the context of plea bargaining, highlighting the tension between procedural efficiency and fundamental guarantees of a fair trial (Angwin et al., 2016; Katz, 2020). AI can reduce the length of cases and standardize practice, but it brings risks related to voluntarism, transparency, equality of arms, and accountability (Goodman & Flaxman, 2017; Floridi, 2021). Predictive models that estimate the probability of conviction or suggest sentences can influence the decision-making autonomy of the defendant (Wachter & Mittelstadt, 2019). The main proposal is to establish a regulatory framework focused on meaningful human control, algorithmic audit and data protection (CEPEJ, 2021), so that the final decision remains a responsible and fair human act. Conclusion: AI can support criminal justice only if it is used transparently, ethically and under constant legal supervision (Dignum, 2022).

Key words: artificial intelligence, voluntarism, transparency, equality of arms .

INTRODUCTION

Context of modernization and digitalization. Over the last decade, judicial systems have experienced an accelerated process of digitalization, driven by the global need for efficiency, cost reduction and expansion of access to justice

(*Sartor & Lagioia, 2020*). In this landscape, AI-based technologies have gone beyond the stage of auxiliary tools to become components intended to directly influence decision-making processes, evidence analysis and drafting of procedural documents (*Surden, 2020*). In criminal justice, AI roles extend beyond support applications — to models capable of assessing risks or probabilities of recidivism, analyzing the coherence of evidence, or suggesting defense strategies (*Katz, 2020*).

The emergence and adoption of technology continues under multiple pressures: shortage of human and material resources, large volumes of files, demands for transparency and predictability from society (*Susskind, 2019*). The transformation is not neutral, however; the introduction of AI raises fundamental questions about the place of algorithms and human decision-making in an act of justice (*Floridi, 2021*). The interest of this study focuses on a particular procedural mechanism — the plea agreement — and on how algorithmic assistance can modify, even involuntarily, the relationship between autonomy, responsibility, and efficiency.

Central theme and motivation of the study. The analytical theme aims to determine to what extent AI can influence the negotiation and conclusion of the plea agreement and to what extent this influence is compatible with the fundamental principles of fair criminal procedure (*Alschuler, 2019*). The central reason is the risk of transforming voluntarism into a technical calculation: if an algorithm suggests a punishment or indicates the probability of conviction, the defendant's decision can be hijacked by the perception of the infallibility of the technology (*Mittelstadt, 2022*).

In addition to the direct risk to free consent, problems arise such as unequal access to digital tools, the lack of transparency on the functioning of algorithmic systems or the lack of clear regulations on liability in case of errors (*Goodman & Flaxman, 2017; Wachter & Mittelstadt, 2019*). The study identifies both the benefits — procedural efficiency, uniform decisions, rationalization of resources — and the dangers associated with the use of AI, proposing concrete normative solutions (*Dignum, 2022*).

Conceptual delimitations. To guide the analysis, a series of essential concepts are defined:
• Legal AI: the set of technologies and algorithms that assist legal analysis or drafting (*Sartor & Lagioia, 2020*).
• Algorithmically assisted decision: the situation in which a human operator uses AI results while retaining final responsibility (*European Commission, 2021*).
• Judicial predictive models: the use of machine learning techniques to estimate the probability of judicial outcomes (*Katz, 2020*).
• Judicial autonomy: the capacity of the system's actors to decide independently, without technological determinations *(Floridi, 2021*).

Maintaining human control over the decision and ensuring transparency are fundamental principles in international documents on the use of AI in justice (*CEPEJ, 2021*).

I. THEORETICAL AND LEGISLATIVE FRAMEWORK
I.1. Plea bargaining in Romanian law

In Romanian criminal procedural law, the plea bargaining is regulated by articles 478–488 of the Code of Criminal Procedure (*Neagu, 2018; Udroiu, 2022*). The institution was introduced as a simplified procedure, intended to streamline the criminal system, reduce the duration of trials and relieve the courts (*Damaschin, 2020*). The mechanism borrows elements from Anglo-Saxon plea bargaining, but adapts them to the principles of continental law, preserving the specific procedural guarantees (*Alschuler, 2019*).

The agreement is materialized when the defendant admits the alleged act, accepts the legal classifications and consents to a fixed sentence, agreed with the prosecutor. Subsequently, the court verifies whether the agreement complies with the legal norms and whether there are reasons for non-validation – such as the lack of voluntarism or the disproportionate nature of the sanction (*Volonciu, 2017*). Thus, judicial review prevents the agreement from becoming an arbitrary instrument and ensures respect for fundamental rights (*Neagu, 2018*).

The fundamental principles are:

1. Voluntarism – the free, conscious and unconstrained will of the defendant (*Alschuler, 2019*);
2. Legality – compliance with the legal framework (*Damaschin, 2020*);
3. Proportionality – the balance between the gravity of the act and the sanction (*Volonciu, 2017*);
4. Judicial review – guaranteeing the independent verification of the agreement.

In the absence of adequate review, procedural efficiency can become an end in itself, with the risk of undermining fundamental rights (*Udroiu, 2022*).

I.2. Theoretical and practical opportunities and challenges

The introduction of AI in agreement negotiation and validation implies a major shift in the relationship between the human factor and the algorithm (*Surden, 2020*). AI can support the analysis of evidence, identify relevant elements, compare situations with previous practice and provide estimates of possible outcomes (*Katz, 2020*). This provides the parties with a more complex and efficient decision-making matrix.

At the same time, the use of AI raises important issues of accountability — who is responsible for decisions influenced by algorithmic recommendations? (*Dignum, 2022*). Difficulties related to transparency and fairness are added: how can one verify that the training data does not contain systemic biases? (*Goodman & Flaxman, 2017*). Risks related to informed consent and equality of arms also

arise in the context where access to advanced tools is not uniform (*Susskind, 2019*).

I.3. Relevant international experiences

Different jurisdictions have approached AI in criminal law with different nuances:

• United States: Plea bargaining is the predominant way to resolve criminal cases (*Alschuler, 2019*). Programs such as COMPAS are used to assess the risk of recidivism, but critical analyses indicate racial and socio-economic biases, affecting the fairness and transparency of the process (*Angwin et al., 2016; Mittelstadt, 2022*).

• United Kingdom: Cautious approach, using AI mainly in administrative phases and for decision-making support, maintaining strict human control to prevent serious errors (*Sartor & Lagioia, 2020*).

• Estonia: Experimenting with a more advanced level of legal automation, including ideas for algorithmic judges for minor cases, but retaining the possibility of human appeal (*Surden, 2020*).

These experiences highlight the need to maintain a balance between integrating technology and preserving the human factor as a guarantor of procedural fairness.

I.4. European guidelines on the use of AI in justice

European documents emphasize accountability, transparency, data protection and meaningful human control. The CEPEJ — a body of the Council of Europe — has analyzed the use of AI in judicial systems and the impact of big data on legal decisions, emphasizing that algorithms must be used in an ethical, controllable and explainable manner (*CEPEJ, 2021*).

These guidelines reflect the concern to guarantee the balance between efficiency and respect for fundamental rights, given that data and algorithms can decisively influence judicial outcomes and the lives of individuals (*European Commission, 2021; Floridi, 2021*).

II. AI APPLICATIONS IN CRIMINAL JUSTICE
II.1. Directions of use

AI in criminal law can be thought of on various levels, in line with international trends regarding the digitalization of justice (*Surden, 2020; Susskind, 2019*):

1. Automated legal assistance — systems that help draft, verify, and analyze legal documents. These can reduce material errors and suggest models, while the lawyer remains responsible for the content (*Susskind, 2019*).

2. Predictive analytics — data from previous cases is used to estimate values or probabilities, supporting the formation of a strategy or risk assessment. The use of these tools derives from the development of quantitative prediction methods in the legal field (*Katz, 2020*).

THE IMPACT OF ARTIFICIAL INTELLIGENCE IN THE CONCLUSION OF THE PLEA AGREEMENT

3. Sentencing recommendations — AI can suggest a punishment range depending on the severity of the offense, the defendant's profile, and previous practice, replicating predictive models known in recent literature *(Sartor & Lagioia, 2020).*

4. Case classification and management — tools that optimize case distribution, identify priorities, reduce waiting times, and allow the generation of useful statistical analyses for resource allocation *(Dignum, 2022).*

II.2. Concrete examples and their impact

• COMPAS — a tool used in the USA to assess the risk of recidivism. It has raised criticism regarding bias and lack of transparency, being extensively analyzed in academic reports and studies (*Angwin et al., 2016; Goodman & Flaxman, 2017*). Its impact emphasizes the need for continuous verification of the correctness and fairness of models *(Mittelstadt, 2022).*

• Prometea — in Argentina, this system automates repetitive procedures and drafts administrative documents, significantly reducing processing time and contributing to the standardization of administrative flows *(Sartor & Lagioia, 2020).*

• ROSS Intelligence — a natural language processing platform used for legal research and argument development. Examples show that, if used correctly, such tools can increase the speed and quality of legal work, without replacing human critical evaluation *(Surden, 2020; Susskind, 2019).*

II.3. Operational advantages and risks

Advantages:

• Increased efficiency — reduced analysis and drafting time, in line with the literature on the digital transformation of legal practice (*Susskind, 2019).*

• Uniformized practice — identifying decision patterns and limiting subjective variations between similar decisions, an aspect emphasized in studies on legal prediction *(Katz, 2020).*

• Increased objectivity — reduced emotional influences or individual biases, to the extent that the data is clean and the algorithms are designed responsibly (*Floridi, 2021; Dignum, 2022*).

• Widely accessible analytics — the ability to quickly generate statistical reports and compare similar situations, facilitating institutional transparency (*European Commission, 2021*).

Risks:

• Algorithmic bias — if the training data reflects historical inequalities, the system may reproduce or amplify such biases *(Angwin et al., 2016; Goodman & Flaxman, 2017).*

• Decisional opacity — the difficulty of challenging an algorithmic recommendation in the absence of access to how the system works, a central issue in discussions about explainability *(Wachter & Mittelstadt, 2019).*

• Technology dependency — the risk that legal actors become excessively dependent on the systems' suggestions, losing the capacity for critical analysis *(Floridi, 2021)*.

• Unequal arms — asymmetric access to advanced tools can create an unjustified procedural advantage, affecting the principle of fair trial *(Sartor & Lagioia, 2020)*.

These risks justify the need for strict regulations on algorithmic transparency, system auditing, legal accountability, and equal access to technology (*CEPEJ, 2021; European Commission, 2021*).

III. THE INCIDENCE OF IA IN THE PLEA AGREEMENT
III.1. Possible roles of AI in negotiation and settlement

In settlement negotiation, AI can intervene at four levels, each with benefits and risks, as evidenced by the legal and ethical literature on the use of digital technologies in justice *(Surden, 2020; Dignum, 2022)*.

III.1.1. Evidence analysis and estimation of the probability of conviction

• AI can process large amounts of data and identify correlations between the described act and previous decisions, providing an approximation of the degree of certainty of guilt (*Katz, 2020)*.

• It can help the parties understand whether the case is solid, if there are major risks of failure, or if there are elements that justify a negotiated solution *(Sartor & Lagioia, 2020)*.

• Major risk: overconfidence in the algorithmic result considered objective, even if the data contains errors or biases or does not reflect legislative or contextual changes *(Mittelstadt, 2022; Goodman & Flaxman, 2017)*.

III.1.2. Formulating plea bargain proposals

• The algorithm can calculate scenarios for various sentences, checking proportionality to the committed act, precedents and legal profiles *(Sartor & Lagioia, 2020)*.

• It can suggest compromise options, so that the parties save time and resources, transforming the negotiation into a more structured process *(Susskind, 2019)*.

• Problems arise from algorithmic opacity: if the defendant does not understand the parameters that determine the outcome, consent is no longer fully informed *(Goodman & Flaxman, 2017; Wachter & Mittelstadt, 2019)*. Lack of transparency can hinder verification of correctness and create a procedural imbalance.

III.1.3. Simulating sentencing scenarios and legal effects

• AI can estimate the length of the sentence, the probability of suspension, the effects on the criminal record and on civic or social rights, aspects analyzed in the literature on legal prediction *(Katz, 2020)*.

• This information helps to assess the proportionality and impact of sanctions.

• The main danger: excessive standardization and rigidity of the decision, reducing the place for the individual peculiarities of the person and the context *(Floridi, 2021; Dignum, 2022)*.

IV.1.4. Generating predictive reports on the sanction

• The analysis of variables such as age, antecedents, type of offense, procedural attitude or socioeconomic conditions can generate useful predictions for procedural actors *(Katz, 2020)*.

• Such reports can be used by the prosecutor, court or defense for substantiation *(Sartor & Lagioia, 2020)*.

• The central question of responsibility — who is responsible for an erroneous result? the developer, the institution or the magistrate who used the recommendation? — is present in the literature on algorithmic liability (Dignum, 2022; CEPEJ, 2021). The need for a clear framework becomes essential to avoid major harm and erroneous decisions.

III.2 Legal and moral dilemmas

1. Validity of consent: if the decision is influenced by an algorithmic suggestion perceived as objective, not by personal conviction, the integrity of voluntarism is affected — an issue intensely discussed in AI ethics *(Floridi, 2021; Dignum, 2022)*.

2. Fairness of the procedure: unequal access to technologies can create major imbalances between parties, affecting equality of arms, a fundamental principle in doctrine and jurisprudence *(CEPEJ, 2021; Susskind, 2019)*.

3. Real expression of will vs. product of a technical calculation: if the agreement is partly generated technically, it risks reflecting an algorithmic rationality focused on efficiency, not on fairness or human responsibility *(Floridi, 2021)*.

III.3. Impact on fundamental rights

– Freedom of expression of consent: subtle pressures generated by statistics and predictions can hijack free decision-making, a phenomenon analyzed in the literature on the influence of AI on autonomy *(Mittelstadt, 2022)*.

– Data confidentiality: the processing and storage of sensitive criminal information by algorithms raises important data protection risks *(European Commission, 2021; Wachter & Mittelstadt, 2019*).

– Impartiality: historical biases can influence algorithmic results, affecting the principles of equality and non-discrimination *(Angwin et al., 2016; Goodman & Flaxman, 2017)*.

– Professional and legal liability: the literature highlights the need for a detailed legal framework on liability in the context of AI-assisted decisions *(Dignum, 2022; CEPEJ, 2021)*.

IV. CRITICAL ANALYSIS AND ETHICAL PERSPECTIVES
IV.1. Influence on decision-making autonomy

AI can exert a subtle form of psychological pressure: people tend to attribute scientific and objective authority to automatic computation, a phenomenon well documented in the literature on the impact of algorithms on decision-making *(Mittelstadt, 2022; Floridi, 2021)*. In situations where there are difficulties in understanding how the system works, an algorithmic recommendation can be accepted without the necessary critical analysis *(Goodman & Flaxman, 2017)*.

This means that the decision is no longer just a personal act of responsibility, but may reflect the level of trust given to the technology, not one's own reasoning *(Dignum, 2022)*. Undermining voluntarism affects the legitimacy of the agreement: if there is no clarity and genuine ownership, the agreement may be contestable.

IV.2. Equality of arms and fair access to technology

The right to a fair trial presupposes that neither party benefits from an unjustified advantage. Technological differences can generate significant imbalances between the prosecution and the defense, depending on who has access to high-performance AI systems *(Susskind, 2019)*.

Without a clear regulatory framework on the access and use of AI, a procedural imbalance is created that can seriously affect the adversarial nature of the criminal trial, a fundamental principle in European jurisprudence *(CEPEJ, 2021)*.

IV.3. Meaningful human control

The principle of meaningful human control requires that any final decision be validated by a responsible person, who understands and critically evaluates how the algorithm reached its conclusions *(Dignum, 2022; Floridi, 2021)*. Control is not a simple formality, but involves:
• examining the criteria used by the algorithm,
• verifying the results,
• correcting potential errors,
• assessing the limits of the system *(Surden, 2020)*.

The absence of this control would transform legal actors into executors of technological recommendations. Adequate professional training and continuous monitoring are essential conditions for compliance with the principle *(European Commission, 2021)*.

IV.4. Professional and ethical responsibility

For professionals involved in criminal proceedings, responsibilities extend to:
– Transparency: disclosing how AI is used, together with its limits and risks *(Goodman & Flaxman, 2017)*.

– Confidentiality: protecting data, in accordance with European regulations on AI and data protection *(European Commission, 2021)*.

– Informed consent: explaining the role of the algorithm and its impact in a way that the parties understand *(Mittelstadt, 2022)*.

– Critical review: prohibiting the substitution of the mechanical interpretation provided by AI for legal analysis *(Floridi, 2021)*.

– Conformity assessment: verifying compliance with non-discrimination and data protection standards *(CEPEJ, 2021)*.

These obligations are directly linked to professional ethics, which today extends to the responsibility of managing technologies in legal activity *(Dignum, 2022)*.

IV.5. Philosophical and moral dimension

Justice is not just the result of a calculation, but a process of interpretation, responsibility and empathy, values that are difficult to translate into algorithmic logic *(Floridi, 2021)*. Excessive delegation to non-human systems raises fundamental questions:

• How much reasoning can be delegated?

• What are the moral consequences of exclusive trust in algorithms?

• Can a system understand human complexity, essential in individualizing punishment?

If AI decides based on a set of data without perceiving the human dimension of the case — suffering, motivations, potential for rehabilitation — the essence of justice is lost *(Mittelstadt, 2022)*. AI can support reasoning, but it cannot replace moral judgment or human responsibility *(Floridi, 2021; Dignum, 2022)*.

The ethics of using AI in justice requires a holistic approach, in which technology is a tool, not a substitute for human discernment.

CONCLUSION

Main conclusions

1. AI can provide valuable support in criminal justice, including in the negotiation of plea agreements, through efficiency, uniformity and increased data analysis capacity (Susskind, 2019; Surden, 2020).

2. The risks are real and serious: impairment of voluntarism, lack of transparency, biases, inequality of access to technology and unclear responsibility in case of errors, aspects highlighted in the critical literature on AI (Mittelstadt, 2022; Angwin et al., 2016; Goodman & Flaxman, 2017).

3. The description of the agreement remains a human act; AI cannot and should not become the final decision-maker, as moral and legal responsibility belong to the human factor (Floridi, 2021; Dignum, 2022).

4. The principle of meaningful human control is essential, being recognized both in AI ethics and in European regulatory documents (European Commission, 2021; CEPEJ, 2021).

5. Transparency, audit and protection of sensitive data are fundamental for the legitimate use of AI in justice, to allow for verification of results and prevention of abuse (Wachter & Mittelstadt, 2019; CEPEJ, 2021).

Recommendations for regulation and practice

1. Clear regulatory framework defining the role of AI in criminal proceedings, the limits of use and the responsibilities of actors, in line with international recommendations on trustworthy AI (European Commission, 2021).

2. Transparency standard: operators must disclose the criteria, data and logic used by algorithms, to allow for a real verification of the decision-making process (Goodman & Flaxman, 2017).

3. Audit and periodic verification: Independent systems should test the accuracy, impartiality and ethical compliance of AI used in justice (CEPEJ, 2021; Dignum, 2022).

4. Professional training: Robust training programmes for magistrates, prosecutors and lawyers are essential to avoid passive automation and encourage critical assessment (Surden, 2020).

5. Fair access mechanisms: To maintain equality of arms, all parties should have a comparable level of access to technology or procedural compensation (Susskind, 2019).

6. Liability regulations: Clarifying liability in the event of algorithmic errors — developer, institution or human operator — is an urgent legal need (Wachter & Mittelstadt, 2019; Dignum, 2022).

7. Robust data protection: the use of AI must be compatible with European data protection requirements and the risks associated with the processing of sensitive information (European Commission, 2021).

Future directions for research and application

– The impact of AI on the professional ethics of magistrates and lawyers, an analysis increasingly present in the literature dedicated to the digitalization of justice (Floridi, 2021; Dignum, 2022).

– Public perception of digitalized justice, as citizens' trust is an essential indicator of the legitimacy of the system (CEPEJ, 2021).

– Good practice guides that include technical, legal and ethical standards, integrated into the continuous training of professionals (European Commission, 2021).

– International comparative assessments to identify effective models and adapt local solutions (Sartor & Lagioia, 2020).

Final Message

The future of plea bargaining in the digital age depends not only on technological developments, but also on the ability of the legal system to maintain

the balance between innovation and the protection of human dignity. Even if AI can bring powerful tools, their use must be foreseen, controlled, transparent and subject to human oversight (Floridi, 2021; Dignum, 2022).

Only in this way can an authentic, credible and dignified criminal justice system be built, in which technology supports, and does not subordinate, human reason and responsibility.

BIBLIOGRAFIE

1. Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks.* ProPublica;
2. CEPEJ – European Commission for the Efficiency of Justice. (2021). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment.* Council of Europe;
3. Damaschin, M. (2020). *Procedura penală. Partea generală.* Editura Universul Juridic;
4. Dignum, V. (2022). *Responsible Artificial Intelligence: Developing and Using AI in a Responsible Way.* Springer;
5. European Commission. (2021). *Ethics Guidelines for Trustworthy AI.* Publications Office of the European Union;
6. Floridi, L. (2021). *The Ethics of Artificial Intelligence.* Oxford University Press;
7. Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a right to explanation. *AI Magazine, 38*(3) 50–57;
8. Katz, D. M. (2020). Quantitative legal prediction. *Emory Law Journal, 62*(4), 909–966;
9. Kelleher, J. D. (2019). *Deep Learning.* MIT Press;
10. Lagioia, F., & Sartor, G. (2020). *The Impact of Artificial Intelligence on the Judiciary.* Springer
11. Mittelstadt, B. (2022). Principles for transparency and accountability in algorithmic decision-making. *AI & Society, 37*(2), 1–15;
12. Neagu, I. (2018). *Tratat de procedură penală. Partea generală.* Editura Universul Juridic;
13. Susskind, R. (2019). *Tomorrow's Lawyers: An Introduction to Your Future* (2nd ed.). Oxford University Press;
14. Surden, H. (2020). Artificial Intelligence and Law: An overview. *Georgia State University Law Review, 35*(4), 1305–1340;
15. Udroiu, M. (2022). *Procedură penală. Partea generală. Sinteze şi grile.* Editura C.H. Beck;
16. Volonciu, N. (2017). *Tratat de procedură penală. Partea generală.* Editura Paideia;

17. Wachter, S., & Mittelstadt, B. (2019). A right to reasonable inferences: Rethinking data protection law in the age of Big Data and AI. *Columbia Business Law Review, 2019*(2), 494–620.