REPORT 6

Artificial Intelligence in the Judiciary from a Speculative Perspective

[ANONYMIZED REPORT – NO IDENTIFYING INFORMATION]

1. Introduction

The use of advanced technologies, such as Artificial Intelligence (AI), in the judiciary has aroused growing academic and practical interest. The search for a more efficient and fair judicial system drives the exploration of new methods and tools capable of improving the administration of justice. In this context, the application of AI emerges as a promising proposal, bringing with it the need for a theoretical deepening that connects its practical applications to the foundations of Information Systems and to a business-oriented perspective.

The present study is based on three foundations. The first consists of the non-trivial correlation between AI and Information Systems in the judicial context, since this requires a deep understanding of business rules and the theoretical models that support the integration of these technologies into judicial processes. The business perspective offers a structured framework to analyze how AI can be implemented efficiently, while ensuring that the fundamental principles of justice are preserved.

Judicial processes, when viewed from a business perspective, reveal a series of rules and procedures that can be optimized through the application of advanced technologies. AI, in particular, has the potential to automate repetitive tasks, analyze large volumes of data, and provide valuable insights that can accelerate case resolution. However, the implementation of AI in the judiciary requires a solid theoretical foundation that considers the ethical and systemic aspects involved.

Subsequently, the study seeks to discuss AI through the lens of systems thinkers, offering an overview of their contributions to the robustness of Information Systems, especially when addressing the relationship between humans and machines. The notion of mechanistic thinking proposed by these theorists is

crucial to understanding how AI can be harmoniously integrated into the judiciary. This perspective allows the establishment of complementarity between human capabilities and the advantages offered by automation, resulting in a more effective and fair system.

Third and finally, the study seeks to discuss the feedback mechanism in the application of AI in the judiciary. The use of continuous feedback ensures that automated decisions are constantly evaluated and corrected when necessary. This process of continuous monitoring and adjustment is fundamental to ensuring that AI operates in accordance with the principles of justice and that any deviations are promptly identified and corrected.

The continuous evaluation of AI decisions by humans is essential to mitigate potential risks associated with automation in the judiciary. Human presence in the feedback process not only increases decision accuracy but also reinforces accountability and transparency, fundamental elements for public trust in the judicial system. The interaction between humans and AI must therefore be carefully structured to maximize the benefits of technology while minimizing risks.

The research seeks to highlight the importance of a robust and well-implemented feedback system. An effective feedback mechanism ensures that the information processed by AI is correctly absorbed and that decisions are adjusted as necessary. This approach not only increases the efficiency of the judicial system but also promotes justice by ensuring that automated decisions are fair and accurate.

2. Methodologies Used

2.1. The Digital Ethics Compass

The first artifact used is *The Digital Ethics Compass*, authored by the Danish Design Center. It is presented by the organization as a tool to help companies make correct decisions from an ethical design perspective. Its use therefore consists of applying design techniques to facilitate the association between

activities and mitigating or complementary actions related to a given action.

In this sense, the Danish Design Center describes that, when using action 04 of the "automation" tab — namely the action: "Is your automated system transparent, allowing the user to see how it works?" (free translation from English: *Is your automated system transparent, so the user can see the engine room?*) — we encounter the following instructions (Danish Design Center, Automation 4):

- Try to explain to the public how your automated systems make decisions.
- Make your algorithms open and accessible so that experts can "open the hood."
- Invite experts to review your algorithms.
- Try to develop simple algorithms with logic understandable to humans.
- As much as possible, avoid black-box systems where you do not even understand how the algorithms work.

These instructions therefore aim to indicate mitigating contours for potential problems experienced by technologies that use that action.

In the context of the present study, the use of an ethical tool is paramount. When discussing the legality and adequacy of AI tools in the judiciary, ethical and moral aspects cannot be dissociated from their use. This notion becomes clearer when analyzing that the use of AI in the judiciary is still recent and embryonic, being received with some resistance and caution (Zalnieriute, 2021, p. 3).

This behavior can, to some extent, be justified by the need to incorporate legal principles as a fundamental part of AI architecture (Katyal, 2022, p. 323); otherwise, rejection of its use will continue to grow, especially due to concerns about its outcomes (Xenidis and Senden, 2020; Liu and Du, 2021).

However, only the *Data* and *Automation* dimensions will be used, since they are more closely related to the case studied. Design analysis, in its strict sense, is not opportune, considering

that graphical implementation is rarely relevant in the AI process.

2.2. Cone of Possibilities

The final tool used in the present study is the *cone of possibilities*. Unlike the others, the cone of possibilities aims to present four distinct scenarios: (a) possible futures; (b) desired futures; (c) probable futures; and (d) plausible futures (Delve, n.d.).

The use of this artifact aims, in a complementary manner to the other tools, to provide a broader and more directed view of the realities that may be perceived from the application of the tool in the real world. Moreover, its inclusion allows a more objective analysis of expected behaviors in these four dimensions of occurrence.

Considering its use, Section 4 will present the third (c) and fourth (d) dimensions, discussing, respectively, the probable and plausible futures of tool usage; while Section 5 will discuss the first (a) and second (b) dimensions, addressing the possible and desired futures of its use.

3. Where Are We? – A Contextualization from the State of the Art

3.1. Thematic Context

The use of advanced technologies, such as Artificial Intelligence (AI), in the judiciary has aroused growing academic and practical interest. The search for a more efficient and fair judicial system drives the exploration of new methods and tools capable of improving the administration of justice. In this context, the application of AI emerges as a promising proposal, bringing with it the need for theoretical deepening that connects its practical applications to the foundations of Information Systems and to a business-oriented perspective.

The present study has three foundations. The first consists of the non-trivial correlation between AI and Information Systems in the judicial context, since this requires a deep understanding of business rules and theoretical models that

support the integration of these technologies into judicial processes. The business perspective offers a structured framework to analyze how AI can be implemented efficiently, while ensuring that the fundamental principles of justice are preserved.

Judicial processes, when viewed from a business perspective, reveal a series of rules and procedures that can be optimized through the application of advanced technologies. AI, in particular, has the potential to automate repetitive tasks, analyze large volumes of data, and provide valuable insights that can accelerate case resolution. However, the implementation of AI in the judiciary requires a solid theoretical foundation that considers the ethical and systemic aspects involved.

Subsequently, the study seeks to discuss AI through the lens of systems thinkers, offering an overview of their contribution to the robustness of Information Systems, especially when addressing the relationship between humans and machines. The mechanistic perspective proposed by these theorists is crucial to understanding how AI can be harmoniously integrated into the judiciary. This perspective allows the establishment of complementarity between human capabilities and the advantages offered by automation, resulting in a more effective and fair system.

Third and finally, the study seeks to discuss the feedback mechanism in the application of AI in the judiciary. The use of continuous feedback ensures that automated decisions are constantly evaluated and corrected when necessary. This process of continuous monitoring and adjustment is fundamental to ensuring that AI operates in accordance with the principles of justice and that any deviations are promptly identified and corrected.

The continuous evaluation of AI decisions by humans is essential to mitigate potential risks associated with automation in the judiciary. Human presence in the feedback process not only increases decision accuracy but also reinforces accountability and transparency, fundamental elements for public trust in the judicial system. The interaction between humans and AI must therefore be carefully structured to maximize the benefits of technology while minimizing risks.

3.2. Mapping of the Current Scenario Surrounding the Topic

Currently, AI has been cautiously employed by the judiciary of several countries. Its use ranges from organizing procedural records, as discussed in Fortes (2020) and Nowotko (2021), to full application throughout the entire process, dispensing with human intervention, as discussed in Katyal (2022) and Kumar et al. (2023).

The mapping is discussed based on Figure 1, which clearly demonstrates the relationship between society, law, and technology.

Society
- Social justice
- Equity
- Moral judgment
- Popular legitimacy

Technology
- Innovation
- Practicality
- Transformation of relationships

Law
- Legal norms
- Equality
- Procedural rules
- Tradition

Figure 1. Intersection between Law and Society.

In the first axis, the notion of law itself is discussed, understood as an autonomous field of knowledge and consolidated as an applied social science. In this sense, it is possible to verify that both norms and operators — those who routinely apply norms in judicial activity — are confronted with the new technological reality.

It is relevant to highlight that for lawyers and legal firms, jurimetrics has already been consolidated. Jurimetrics is a branch of law that seeks, through the reading of patterns and trends, to indicate which cases have better or worse outcomes,

optimizing time and the decision-making process regarding case acceptance.

These jurimetric tools are not available to judges, since it is not within their competence to choose which cases to accept. Moreover, jurimetrics does not aim to indicate steps or procedures to be applied in the process, but merely to provide statistical data to lawyers, leaving them in charge of conducting the entire process.

The context observed in the review indicates a paradigm shift, such that technology ceases to play a purely secondary role and begins, at minimum, to act in parallel with the professional. This shift certainly implies conceiving a new configuration of laws and training, to ensure its correct use.

In the social axis, it is possible to observe its relevant and indispensable condition of legitimizing the use of AI within the judicial context. As discussed by Das et al. (2019), Fortes (2020), and Al-Obeidi and Al-Mulla (2022), for a decision to be socially recognized as valid, the entire group must recognize the legitimacy of its source. That is, for AI to have the power to issue a valid and enforceable decision, it must be anchored in basic legal principles, legitimizing its enforcement before society and demonstrating itself capable of judging as well as — or better than — a human judge.

Thus, society is an essential pillar for AI application in the judiciary. Without it, use would not only be weakened in its foundations but would also lack elements that give robustness to the resulting decision.

Finally, the technological axis demonstrates the fragmented use of AI in law. This axis encompasses both the use of AI itself — through processes of organizing procedural records and conducting the entire procedural rite — and limitations of its use, such as biases and constraints related to the size of legal documents.

It is relevant to observe that the technological axis always permeates the other two axes, since society continuously develops and law, especially civil law, seeks to develop norms aimed at regulating and protecting society from potential harms caused by technology use. Law, however, is not immune to social

changes, especially technological ones. Examples include the use of computers in judicial courts, virtual hearings, and the digitization of physical cases, optimizing judicial activity.

It is therefore possible to infer that AI begins to influence law, as it presents itself as a tool capable of solving daily problems faced by the judiciary. However, since it affects an indeterminate number of people and presupposes access to confidential information of the parties, it must be incorporated gradually, with constant verification of its adequacy.

4. Where Are We Going? – A View of Possible Futures of the Technology

4.1. Temporal Projection

Despite advances seen in Eastern judiciaries (Nowotko, 2021; Papagianneas and Junius, 2023), it is possible to state that in the West this advance is slower, especially due to normative configuration. Moreover, in the national context (Fortes, 2020; Silva and Costa-Abreu, 2022), this use is still embryonic and applied only in a punctual manner by courts.

Considering this reality, a period of ten years is estimated as a reasonable timeframe for a new assessment of the reality of its use.

4.2. Mapping the Future Scenario Based on Trends

Following the discussion in Section 3.2, three positive and two negative trends are observed. The positive trends are: (a) use of AI to organize procedural records; (b) use of AI as a tool to assist the judge (human judge); and (c) complete replacement of the judge by the machine (AI). The negative trends consist of: (a) data integrity; and (b) existence of input biases, known and unknown.

The present discussion is based on the Danish Design Center, establishing a relationship with the automation and data axes. The discussions are based on the permanence of behaviors observed in the review and discussed in previous sections, that is, assuming the continuation of the same forces already identified, ignoring changes in this course.

The first discussion is based on the compass question: “Are users aware that they are interacting with an automated solution?” This concern is fundamental and also discussed in the literature, albeit subtly and with little emphasis (Fortes, 2020; Al-Obeidi and Al-Mulla, 2022; Petkovic, 2023). This discussion is particularly relevant in the European Union and Brazil, due to the GDPR and LGPD, respectively.

The compass seeks to prevent users — in this case litigants — from interacting with AI without knowledge of its presence. According to the legal panorama, both GDPR and LGPD provide for the possibility of reviewing automated decisions, requiring that such decisions be reviewed by a human.

In the hypothesis of complete replacement of the judge by AI, it is necessary that this substitution be publicly disclosed and that legal instruments be guaranteed for full case review, not merely decision review, since the latter may be tainted by incurable defects.

Other compass dimensions discuss whether the automated system complies with legislation and human rights and whether automation is sufficiently transparent. These questions are central to AI use in the judiciary. If AI can violate national legislation or human rights, its use becomes problematic. Therefore, in the case of judge substitution, constant verification mechanisms are necessary to ensure compliance.

The question of whether automation causes people to lose the ability to perform activities is answered negatively, since even in the event of substitution, highly skilled professionals will still be required to train AI, review its decisions, and identify violations.

Bias is another central issue. Algorithms are not biased per se, but training data may lead to undesirable outcomes. Thus, review procedures must be established before AI decisions produce legal effects.

Finally, adaptability to legal changes is discussed. For AI to be correctly applied in the judiciary, it must be capable of learning new laws and applying them correctly. However, the literature indicates slow and inflexible training processes,

suggesting that AI may represent a dangerous failure in judicial service provision if improperly implemented.

5. Where Do We Want to Go? – Desired Future

5.1. Designing an IT Solution

Considering the current scenario, the appropriate solution is to move toward the plausible future, reducing AI's role and allocating it as an auxiliary rather than substitutive tool for human action. This paradigm shift mitigates negative effects and allows constant monitoring of AI evolution.

The proposed information technology solution would establish a link between the judge and technology, developing a feedback and validation system that allows the former to always modify the results of the latter, training and adjusting it for better performance.

5.2. Complementary Actions

Considering the extensions of AI in the judiciary and its effects, the need for gatekeepers is estimated to mitigate damages caused by the system.

The first necessary gatekeeper is professional training, enabling identification of biases and damages caused by the machine and promoting corrections.

The second gatekeeper is the development of guidelines, both legal and institutional, to indicate limits of AI action.

The development of these gatekeepers is essential to ensure that AI remains within its original limits, avoiding indiscriminate use.

6. Conclusion

Throughout the work, the use of AI in the judiciary was discussed from a speculative perspective, based on the literature review.

The probable future presents a dangerous characteristic: low attention to mitigating existing input biases. Given that AI is

only as good as its training data, decisions may be biased depending on the characteristics of the parties and the process.

The probable future reduces judicial autonomy, preventing judges from improving or correcting AI.

The preferred future is utopian and hypothetical, marked by the absence of biases.

The plausible future, less optimistic, limits AI use to punctual and assistive roles, distant from the decision-making process.

This is understood as the best form of AI action, as it preserves human validation capable of correcting errors.

The broad and indiscriminate use of AI in the judiciary may endanger judicial autonomy and must be avoided until there is clarity regarding its “thinking” processes and elimination of biases.

References

(References preserved verbatim from the original document; no changes applied.)

Danish Design Center. Toolkit: The digital ethics compass. https://ddc.dk/tools/

toolkit-the-digital-ethics-compass. [Accessed 04-07-2024].

Monika Zalnieriute. Technology and the courts: Artificial intelligence and judicial impartiality.

SSRN Electronic Journal, 2021. ISSN 1556-5068. doi: 10.2139/ssrn.3867901.

URL http://dx.doi.org/10.2139/ssrn.3867901.

Sonia Katyal. Democracy and distrust in an era of artificial intelligence. Journal of the

American Academy of Arts & Sciences, pages 1–13, 2022. URL https://ssrn.

com/abstract=4099147.

Rapha¨ele Xenidis and Linda Senden. Eu non-discrimination law inthe era of artificial

intelligence: Mapping the challenges of algorithmic discrimination. General Principles

of EU law and the EU Digital Order, 2020. URL https://papers.ssrn.com/

sol3/papers.cfm?abstract_id=3529524.

Peng Liu and Yong Du. Blame attribution asymmetry in human–automation cooperation.

Risk Analysis, 42(8):1769–1783, 2021. ISSN 1539-6924. doi: 10.1111/risa.13674. URL

http://dx.doi.org/10.1111/risa.13674.

Delve. Speculative Design and a Cone of Possibilities

— delve.com. https://www.delve.com/insights/

speculative-design-and-a-cone-of-possibilities, s.d. [Accessed

04-07-2024].

Rubim Borges P. Fortes. Paths to digital justice: Judicial robots, algorithmic decisionmaking,

and due process. ASIAN JOURNAL OF LAW AND SOCIETY, 7(3):453–469,

2020. ISSN 2052-9015 2052-9023. doi: 10.1017/als.2020.12.

Paweł Marcin Nowotko. Ai in judicial application of law and the right to a court. Procedia

Computer Science, 192:2220–2228, 2021. ISSN 1877-0509. doi: https://doi.

org/10.1016/j.procs.2021.08.235. URL https://www.sciencedirect.com/

science/article/pii/S1877050921017324.

S. Kumar, K. Joshi, S. Singh, H. Anandaram, K. S. Sindhu, S. K. Mahariya, and V. Kaushik.

Ai technological interference in court proceedings: Right to fair trial decision. In

2023 3rd International Conference on Advancement in Electronics & Communication

Engineering (AECE), pages 237–242, 2023. doi: 10.1109/AECE59614.2023.10428630.

A. K. Das, A. Ashrafi, and M. Ahmmad. Joint cognition of both human and machine

for predicting criminal punishment in judicial system. In 2019 IEEE 4th International

Conference on Computer and Communication Systems (ICCCS), pages 36–40, 2019.

doi: 10.1109/CCOMS.2019.8821655.

A. H. Al-Obeidi and M. S. Al-Mulla. The legal basis of the right to explanation for artificial

intelligence decisions in uae law. In 2022 International Arab Conference on Information

Technology (ACIT), pages 1–4, 2022. ISBN 2831-4948. doi: 10.1109/ACIT57182.2022.

9994088.

Tania Sourdin, Bin Li, and Donna Marie McNamara. Court innovations and access to

justice in times of crisis. Health Policy and Technology, 9(4):447–453, 2020. ISSN

2211-8837. doi: https://doi.org/10.1016/j.hlpt.2020.08.020. URL https://www.

sciencedirect.com/science/article/pii/S2211883720300927.

B. S. F. Silva and M. Costa-Abreu. Exploring bias analysis on judicial data using machine

learning techniques. In 2022 12th International Conference on Pattern Recognition

Systems (ICPRS), pages 1–7, 2022. doi: 10.1109/ICPRS54038.2022.9854068.

G. Sukanya and J. Priyadarshini. Modified hierarchical-attention network model for legal

judgment predictions. Data & Knowledge Engineering, 147:102203, 2023. ISSN

0169-023X. doi: https://doi.org/10.1016/j.datak.2023.102203. URL https://www.

sciencedirect.com/science/article/pii/S0169023X23000630.

Kalliopi Terzidou. The use of artificial intelligence in the judiciary and its compliance with

the right to a fair trial. Journal of Judicial Administration, 31(154):1–15, 2023. URL

https://ssrn.com/abstract=4495715.

Essoujaa Mohammed, Esghir Mustapha, and Azhari Mourad. Using machine learning

to predict public prosecution judges decisions in moroccan courts. Procedia Computer

Science, 220:998–1002, 2023. ISSN 1877-0509. doi: https://doi.org/10.1016/

j.procs.2023.03.138. URL https://www.sciencedirect.com/science/

article/pii/S1877050923006737.

Giovana Lopes. Artificial intelligence and judicial decision-making: Evaluating the role

of ai in debiasing. TATuP - Zeitschrift f¨ur Technikfolgenabsch¨atzung in Theorie und

Praxis, 33(1):28–33, 2024. ISSN 2568-020X. doi: 10.14512/tatup.33.1.28. URL

http://dx.doi.org/10.14512/tatup.33.1.28.

N. Chugh. Risk assessment tools on trial: Lessons learned for “ethical ai” in the criminal

justice system. In 2021 IEEE International Symposium on Technology and Society

(ISTAS), pages 1–5, 2021. ISBN 2158-3412. doi: 10.1109/ISTAS52410.2021.9629143.

A. J. McLoughney, J. M. Paterson, M. Cheong, and A. Wirth. ‘emerging proxies’ in

information-rich machine learning: a threat to fairness? In 2023 IEEE International

Symposium on Ethics in Engineering, Science, and Technology (ETHICS), pages 1–1,

2023. doi: 10.1109/ETHICS57328.2023.10155045.

A. M. John, U. A. M, and J. T. Panachakel. Ethical challenges of using artificial intelligence

in judiciary. In 2023 IEEE International Conference on Metrology for eXtended Reality,

Artificial Intelligence and Neural Engineering (MetroXRAINE), pages 723–728, 2023.

doi: 10.1109/MetroXRAINE58569.2023.10405688.

D. Petkovic. It is not “accuracy vs. explainability”—we need both for trustworthy ai

systems. IEEE Transactions on Technology and Society, 4(1):46-53, 2023. ISSN

2637-6415. doi: 10.1109/TTS.2023.3239921.

N. Wang. “black box justice”: Robot judges and ai-based judgment processes in china’s

court system. In 2020 IEEE International Symposium on Technology and Society (ISTAS),

pages 58–65, 2020. ISBN 2158-3412. doi: 10.1109/ISTAS50296.2020.9462216.

Sandra Wachter, Brent Mittelstadt, and Chris Russell. Why fairness cannot be automated:

Bridging the gap between eu non-discrimination law and ai. Computer

Law & Security Review, 41:105567, 2021. ISSN 0267-3649. doi: https://doi.

org/10.1016/j.clsr.2021.105567. URL https://www.sciencedirect.com/

science/article/pii/S0267364921000406.

P. K. Lohia, K. Natesan Ramamurthy, M. Bhide, D. Saha, K. R. Varshney, and R. Puri. Bias

mitigation post-processing for individual and group fairness. In ICASSP 2019 - 2019

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),

pages 2847–2851, 2019. ISBN 2379-190X. doi: 10.1109/ICASSP.2019.8682620.

Oleg Metsker, Egor Trofimov, Max Petrov, and Nikolay Butakov. Russian court decisions

data analysis using distributed computing and machine learning to improve lawmaking

and law enforcement. Procedia Computer Science, 156:264–273, 2019. ISSN

1877-0509. doi: https://doi.org/10.1016/j.procs.2019.08.202. URL https://www.

sciencedirect.com/science/article/pii/S1877050919311214.

Felicity Zalnieriute, Monika; Bell. Technology and the judicial role. pages 1–38,

2019. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_

id=3492868.

Emre Mumcuog˘lu, Ceyhun E. O¨ ztu¨rk, Haldun M. Ozaktas, and Aykut Koc¸ . Natural

language processing in law: Prediction of outcomes in the higher courts of turkey.

Information Processing & Management, 58(5):102684, 2021. ISSN 0306-4573. doi:

https://doi.org/10.1016/j.ipm.2021.102684. URL https://www.sciencedirect.

com/science/article/pii/S0306457321001692.

Weslei Gomes Sousa, Rafael Antunes Fidelis, Paulo Henrique Souza Bermejo, Ana Gersica

Silva Gonçalo, and Bruno Souza Melo. Artificial intelligence and speedy trial

in the judiciary: Myth, reality or need? a case study in the brazilian supreme court

(stf). Government Information Quarterly, 39(1):101660, 2022. ISSN 0740-624X. doi:

https://doi.org/10.1016/j.giq.2021.101660. URL https://www.sciencedirect.

com/science/article/pii/S0740624X21000964.

Straton Papagianneas and Nino Junius. Fairness and justice through automation in china’s smart courts. Computer Law & Security Review, 51:105897, 2023. ISSN 0267-3649. doi:

https://doi.org/10.1016/j.clsr.2023.105897. URL https://www.sciencedirect.

com/science/article/pii/S0267364923001073.