Strasbourg, 19 December 2025

CEPEJ(2025)18Final

EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ)

GUIDELINES ON THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE FOR COURTS

Adopted at the 45th plenary meeting of the CEPEJ (Strasbourg, 4 and 5 December 2025).

Contents

Executive Summary 3
1. Introduction 4
1.1. Preamble 4
1.2. Generative AI 4
1.3. Generative AI in the context of the justice sector 5
2. The Guidelines 6
2.1. Operational aspects for generative AI in courts 6
2.1.1. Start with identifying the problem, not the technical solution 6
2.1.2. Identify areas suitable for the use of generative AI 7
2.1.2.1. Case management and judicial administration 7
2.1.2.2. Document management and information processing 7
2.1.2.3. Information services and user assistance 7
2.1.2.4. Drafting support and judicial documentation 7
2.1.3. Be aware of unsuitable applications 7
2.1.4. Prefer customised, secure tools over generic, off the shelf solutions 8
2.1.5. Ensure sound management in the deployment of generative AI systems 10
2.1.6. Be aware of how the parties are using generative AI 11
2.1.7. Ensure awareness and training of AI users 12
2.2. Normative aspects for generative AI in courts 13
2.2.1. Right to a fair trial and judicial independence (Article 6 ECHR) 13
2.2.1.1. Fair trial guarantees 13
2.2.1.2. Principle of legality, legal certainty and predictability 13
2.2.1.3. Exclusivity of judicial authority 14
2.2.1.4. Mandatory human responsibility and supervision 14
2.2.1.5. Judicial independence and non-binding nature of AI 14
2.2.1.6. Transparency and adequate reasoning 14
2.2.1.7. Right of access to a court 15
2.2.2. Right to liberty and security (Article 5 ECHR) 15
2.2.3. Right to privacy and data protection (Article 8 ECHR) 15
2.2.4. Right to an effective remedy (Article 13 ECHR) 16
2.2.5. Prohibition of discrimination (Article 14 ECHR and Protocol No. 12) 16
2.2.5.1. Non-discrimination principle 16
2.2.5.2. Neutrality and active bias prevention 16
2.2.6. Additional Safeguards 16
2.2.6.1. Explainability and transparent reasoning 16
2.2.6.2. Proportionality and caution in generative AI deployment 16
2.2.6.3. State liability for damages caused by generative AI use 17
2.2.6.4. Principle of procedural fairness and efficiency 17
3. GLOSSARY 17

EXECUTIVE SUMMARY

These Guidelines provide a framework for the implementation of generative artificial intelligence (AI) in the administration of justice, with a focus on ensuring legal certainty, judicial independence with effective human oversight, transparency, and traceability.

At the outset, general-purpose tools from large technology companies are only a first step; moving towards specialised solutions and, where appropriate, local Large Language Models (LLMs) or deployments with sovereign control over data and infrastructure. It is essential that tools are tailored to judicial processes and deployed under public control.

The implementation of these solutions shall be in accordance with the European Convention on Human Rights (ECHR), the Council of Europe Framework Convention on Artificial Intelligence (CETS 225), Convention 108+ for the protection of individuals with regard to the processing of personal data and the CEPEJ European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment.

These Guidelines are based on the following core principles: the exercise of judicial power is the exclusive responsibility of the courts; effective access to a human judge is always guaranteed; outputs produced by generative AI are never binding; and the use of generative AI in judicial activities and the drafting of decisions must be transparent. Privacy requirements mandate the pseudonymisation of data and exclusive control of data and infrastructure by the public authority competent for the provision of information technology services to the judiciary. It is vital to consider the legality, judicial independence, non-discrimination, right to an effective remedy and the traceability and explainability of reasoning.

A phased implementation is proposed, with pilots, metrics, and the ability to reverse course if necessary. This approach is underpinned by the principle of subsidiarity, which emphasises the identification of functional needs and the consideration of conventional technological solutions as a priority. Generative AI should be deployed only when alternative solutions are deemed to be ineffective or inefficient.

Its use can be beneficial in a variety of professional contexts, including case management, document processing, public information services and support for drafting routine judicial documents. When used to assist judicial authorities in researching and interpreting facts and law, in applying the law to concrete sets of facts, or in whichever other support to adjudication, compliance with stringent requirements should be followed. To ensure a successful implementation, it is essential to conduct thorough ex-ante risk and impact assessments prior to any deployment. Furthermore, implementing continuous monitoring and feedback mechanisms will enable both the court and the administration of justice to make adjustments and improvements. Finally, training for all court personnel and system users is crucial to ensure the effective use of the system.

Finally, the State should remain liable for any damages arising from the use of generative-AI systems in the administration of justice, ensuring that users retain effective avenues for recourse and remedy in the event of harm.

1. INTRODUCTION

1.1. Preamble

1. These guidelines aim to provide practical advice on how to safely implement generative Artificial Intelligence (AI) in the administration of justice. They are addressed to authorities responsible for the administration of justice and judicial professionals. They should be periodically reviewed and updated in line with recent technological developments.

2. The use of AI in justice shall respect the rights guaranteed by the European Convention on Human Rights (ECHR), the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS225), the protection of personal data in the light of the Council of Europe's Convention for the Protection of Personal Data (CETS108), as well as the principles set out in the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment (the CEPEJ Charter), and where applicable Regulation (EU) 2024/1689 of the European Parliament and of the Council of June 13, 2024 laying down harmonised rules on artificial intelligence (AI Act).

3. These guidelines have been prepared by the CEPEJ Working Group on Cyberjustice and AI (CEPEJ-GT-CYBERJUST) based on the contribution of experts.[1] They consider preliminary work by the CEPEJ such as the "Information note on the use of generative AI by judicial professionals in a work-related context", the "Reflections of the CEPEJ's Artificial Intelligence Advisory Board on how to make Artificial Intelligence work for the judiciary" and the "1st Report on the use of Artificial Intelligence (AI) in the judiciary, based on the information contained in the CEPEJ's Resource Centre on Cyberjustice and AI" and are illustrated through tools presented in the CEPEJ's Resource Centre on Cyberjustice and AI.

1.2. Generative AI

4. Generative AI is a branch of AI that focuses on teaching machines to create new content that convincingly imitates human productions. Instead of merely classifying or predicting data, these systems learn from large datasets to identify underlying patterns and structures, enabling them to generate text, images, music, videos, and other type of content.[2]

5. In the field of text generation, the evolution of "natural language processing" tools has allowed automated software to correctly interpret and generate language naturally. Previously, human language understanding represented a challenge for AI software due to its ambiguity and complexity, but with advances in deep learning, models have become capable of learning the patterns of our language and producing coherent responses or texts. This has meant a paradigm shift in how machines process and generate content from textual content.

6. The technology behind generative AI is based on deep learning and advanced neural network architectures. In essence, these models known as *Large Language Models* or *LLMs*, are trained with vast volumes of data to identify complex patterns (whether word sequences, image pixels, or musical notes) which allows to generate new content that mimics the characteristics of the original material.

7. Despite its effectiveness at a wide range of tasks and adoption in *chatbots*, generative AI currently presents limitations, among which are "hallucinations". In natural language generation, a hallucination is output that the model produces without faithful support in the input or reference knowledge: it can explicitly contradict the source text (intrinsic) or add

[1] Javier Ercilla Garcia, Judge, Spain.
[2] Warankar, M., & Patil, R. (2024). Generative Artificial Intelligence. International Journal of Scientific Research in Engineering and Management, 8(04). doi:10.55041/IJSREM31146

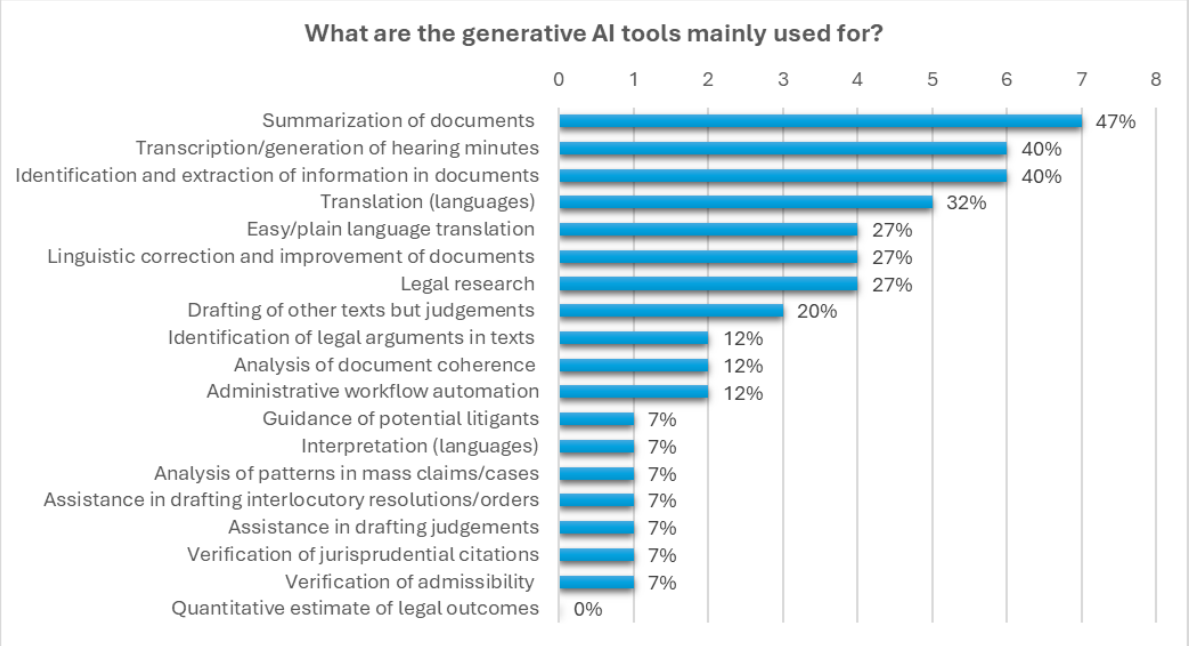
impossible or false data (extrinsic). In essence, it is invented information that the system presents as if it were fact.[3]

8. Other noteworthy limitations concern potential copyright issues arising from the original training data of these models, as well as the unsolicited potential disclosure of input data to third parties and its reuse to train the generative AI model, potentially in breach of applicable data protection regulation.

1.3. Generative AI in the context of the justice sector

9. The use of generative AI solutions by justice professionals is growing, whether through summary generation, text correction, drafting legal foundations based on case-law, interaction with litigants (explaining sentences to clients), etc. According to the *Future Ready Lawyer* 2024 report,[4] which surveyed 712 legal professionals in the United States and nine European countries, 68% of law firms and 76% of corporate legal departments use generative AI at least weekly, with more than a third using it daily (35% in corporate legal departments and 33% in law firms). In Europe, the Netherlands leads daily generative AI use with 42%, followed by Germany at 38%. The primary use of these tools is for document automation (67%), time management (66%), and strategic planning (62%). Furthermore, 58% of law firms plan to increase their AI investment over the next three years, and 73% of corporate legal departments.

10. According to a survey among the CEPEJ's European Cyberjustice Network members from October 2025, 46% of respondents[5] confirmed the use of generative AI in courts of which 74% refer to standard/off the shelf applications (e.g. Microsoft Copilot, Chat GPT, DeepL) and 52% to customised/tailored applications (both options possible). Most mentioned uses of generative AI concern the summarisation of documents and other editorial support functions. No respondent mentioned the use of generative AI for the quantitative estimation of legal outcomes (see Figure 1).



[3] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Chen, D., Dai, W., Chan, H. S., Madotto, A., & Fung, P. (2022). Survey of hallucination in natural language generation. arXiv. https://arxiv.org/abs/2202.03629
[4] Wolters Kluwer. (2024). 2024 Future Ready Lawyer Survey Report: Legal innovation: Seizing the future or falling behind? Wolters Kluwer Legal & Regulatory. https://www.wolterskluwer.com/en/know/future-ready-lawyer-2024
[5] n=33 respondents of 43 addressed Council of Europe member States. Decimals rounded to whole numbers.

Figure 1: Use of generative AI in courts.

11. Various countries are facilitating the use of generative AI tools in public administration and courts through confidentiality agreements, protection of usage data, and guarantees that such data will not be reused for model retraining.[6] [7]

12. In the legal field, the use of generative AI has resulted in lawsuits with false legal citations,[8] [9] which has sometimes led to sanctions for its users.[10] Since "hallucinations" are inevitable[11] or at best mitigable,[12] according to the current state of the art, this is a risk to consider when employing generative AI in justice, without adequate human supervision safeguards.

2. THE GUIDELINES

2.1. Operational aspects for generative AI in courts

2.1.1. Start with identifying the problem, not the technical solution

13. The principle of technological neutrality states that public administrations should focus on functional needs and, on this basis select technological solutions that can be adapted over time. This approach aims to minimise technological dependencies, avoid imposing specific technical implementations or products, and remain adaptable in a rapidly evolving technological environment. This principle should extend to the use of generative AI within judicial administration through a **principle of subsidiarity**, whereby judicial administrations first clearly identify functional needs and second consider different technological solutions, conventional or novel, before opting for the most suitable and least risky. This should not prevent courts from continuously striving to improve efficiency of justice with the support of technology where appropriate, nor from refraining from introducing generative AI in well-established use cases that meet well-identified functional needs.

14. Both courts and administrations of justice face various challenges, including staff shortages, high volumes of mass claims, repetitive drafting tasks, inefficient file management, procedural delays, difficulties in finding jurisprudence, or communication problems with justice system users and lack of interoperability between current IT systems in use.

15. Once operational needs and problems are identified within the administration, the next step is selecting appropriate technical solutions in accordance with the principles of technological neutrality and subsidiarity of AI.

---

[6] Artificial Lawyer. (2025, April 24). UK Courts roll out Microsoft Copilot for judges, update GenAI rules. https://www.artificiallawyer.com/2025/04/24/uk-courts-roll-out-microsoft-copilot-for-judges-update-genai-rules/

[7] El País (2025, 22 September). Do you already know what carpeta.jusitcia.es is? Retrieved from https://cincodias.elpais.com/extras/2025-09-22/ya-sabes-lo-que-es-carpetajusticiaes.html

[8] Weiser, B. (2023, May 27). Here's what happens when your lawyer uses ChatGPT. The New York Times. https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html

[9] MIT Sloan Teaching & Learning Technologies. (2024, November 12). When AI gets it wrong: Addressing AI hallucinations and bias. https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/

[10] Carrick, D., & Kesteven, S. (2023, June 24). This US lawyer used ChatGPT to research a legal brief with embarrassing results. We could all learn from his error. ABC News. https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068

[11] Xu, Z., Jain, S., & Kankanhalli, M. (2024). Hallucination is Inevitable: An Innate Limitation of Large Language Models. arXiv. https://arxiv.org/abs/2401.11817

[12] Nie, F., Yao, J.-G., Wang, J., Pan, R., & Lin, C.-Y. (2019, July). A simple recipe towards reducing hallucination in neural surface realisation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 2673-2679). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1256

16. When selecting the most appropriate technical solutions, cost efficiency should be considered. This should include factors such as energy consumption, cyber security measures, licensing, maintenance, development and consultancy costs, in addition to the initial purchase price.

2.1.2. Identify areas suitable for the use of generative AI[13]

2.1.2.1. Case management and judicial administration

17. Generative AI could perform predictive workload management by analysing historical patterns and create jurisprudence repositories organised by case type and subject matter. These tools could facilitate the identification of argumentative patterns in similar claims, allowing the detection of proceedings that can be resolved uniformly when they raise the same legal question, or helping to allocate cases. Generative AI could also enable the automated recognition of disputed issues by comparing complaints and responses.

2.1.2.2. Document management and information processing

18. Generative AI could automatically generate hearing minutes, draw up thematic and chronological indexes for voluminous case files, and extract relevant data from unstructured documents. Likewise, it could be used to obtain key information, such as names, dates, and amounts, automatically incorporating them into predefined templates.

2.1.2.3. Information services and user assistance

19. Generative AI could allow the creation of specialised chatbots by legal subject matter and conversational procedural-guidance systems for citizens. These applications must be designed with particular attention to the accuracy of the information provided and the clarity of communication with non-specialised users, employing plain language.

2.1.2.4. Drafting support and judicial documentation

20. Generative AI could be used for the automatic creation of standard procedural documents, for the assistance in drafting orders, and the verification of document consistency. It could generate structured summaries of judgments, procedural documents, and appeals, requiring judicial verification that the summary faithfully matches the original. It could also facilitate the preparation of documents, such as reports, outlines, or mind-maps, that present an overview of the points of agreement and disagreement.

21. Furthermore, it could be used for the analysis of complaints and responses with suggestions of related case law; identification and synthesis of the jurisprudence relevant to each case; preliminary analysis of appeals regarding compliance with formal and technical requirements; and linking arguments between different procedural writings, producing organised documents in which the defendant's response is systematically associated with each of the plaintiff's arguments.

2.1.3. Be aware of unsuitable applications

22. The EU AI Act (Regulation (EU) 2024/1689) establishes specific prohibitions on AI practices that are particularly relevant to the administration of justice and potentially generative AI. Article 5 of the Act sets forth the following prohibited AI practices, that can serve as an

[13] The following list is not exhaustive.

illustration also for countries outside its scope of application: AI systems deploying subliminal or manipulative techniques (Article 5(1)(a)); AI systems exploiting vulnerabilities (Article 5(1)(b)); Social scoring systems (Article 5(1)(c); Risk assessment systems for predicting criminal offences (Article 5(1)(d)); Facial recognition database expansion systems (Article 5(1)(e)); Emotion inference systems (Article 5(1)(f)).

23. Additionally, while not explicitly prohibited, the EU AI Act classifies certain AI systems in the justice domain as high-risk under Annex III, particularly those intended to assist judicial authorities in researching and interpreting facts and law, and in applying the law to concrete sets of facts. These systems require compliance with stringent requirements but are not prohibited outright.

24. While LLMs can be useful in the collection of evidence at the investigation stage, they should not replace the judicial assessment of evidence, once introduced in the trial. Although there are strands of logical and neuro-symbolic AI that integrate causal models,[14] generative AI lacks the capacity for *"prospective theorisation and the development of causal logic".*[15] [16]

25. Decision-making should remain under human control. Cases requiring novel normative interpretation or analysis of abstract legal principles should not be entrusted to AI, given that it lacks *"intrinsic understanding of what is true or false in reality".*[17] Likewise, the use of generative AI should be avoided in matters requiring the evolutionary interpretation of recent regulations, given that generative AI may show limitations in adapting its understanding to innovative regulations if it has not been specifically trained with them.[18]

26. The analysis of constitutionality or compliance with international treaties requires a level of axiological interpretation that exceeds the current capabilities of AI. Even if this were not the case in the future, the issue can only be left to human judgment, since it affects the way in which citizens have decided to govern themselves. Therefore, the "zeitgeist" of each moment will influence "constitutional and conventionality control", and this cannot be left to the discretion of an AI.

2.1.4. Prefer customised, secure tools over generic, off the shelf solutions

27. The distinction between generic commercial generative AI tools and customised judicial generative AI is highly relevant for the administration of justice because it entails not merely technical specifications but fundamental issues of data protection, judicial independence, and the integrity of legal proceedings.

28. Free commercial generative AI tools present significant risks. When judicial professionals utilise non-official generative AI services, whether free or paid commercial solutions, they risk inadvertently to incur in an international data transfer, mainly to data centres in the U.S., hosted by the companies providing these services. This data subsequently becomes available for model retraining purposes, effectively removing it from judicial administrative control. Such practices violate fundamental principles of data protection

[14] Colelough, B. C., & Regli, W. (2025). Neuro-symbolic AI in 2024: A systematic review. arXiv. https://arxiv.org/abs/2501.05435
[15] Felin, T., & Holweg, M. (2024). Theory Is All You Need: AI, Human Cognition, and Decision Making. Available at: https://ssrn.com/abstract=4737265.
[16] Li, X., Cai, Z., Wang, S., Yu, K., & Chen, F. (2025). A survey on enhancing causal reasoning ability of large language models. arXiv. https://arxiv.org/abs/2503.09326
[17] Ibid.
[18] Ercilla García, J. (2025). Justicia automatizada: entre las inteligencias artificiales que fingen y las que persuaden. Lex Social, Revista De Derechos Sociales, 15 (1), 1-39. https://doi.org/10.46661/lexsocial.11652

established in Convention 108+[19] and create unacceptable vulnerabilities in the judicial information ecosystem.

29. Generic commercial generative AI solutions, even when offered through paid subscriptions, typically lack the specialised training necessary for legal reasoning and judicial documentation. These tools, designed for general consumer use, demonstrate limited understanding of legal terminology, procedural requirements, and the nuanced reasoning patterns characteristic of judicial work. The absence of integration with verified legal databases increases the risk of "hallucinations", the generation of non-existent legal provisions or misrepresented jurisprudence, which poses severe threats to legal certainty.

30. In case of resorting to commercial solutions, the deployment of generative AI within the administration of justice should be exclusively mediated through formal contractual arrangements between public authorities and AI service providers. In case of resorting to in-house solutions, the deployment of generative AI within the administration of justice should be mediated through a policy on the use of AI in the administration of justice[20] or a broader information security scheme.[21] These contracts or policies should establish comprehensive frameworks addressing:

- Strict data governance protocols ensuring no unauthorised transfer to third parties
- Mandatory data destruction procedures following specified retention periods
- Explicit prohibitions against using judicial data for model retraining without express authorisation
- Clear delineation of liability for data breaches or system failures
- Compliance with national and European data protection regulations
- Regular security audits and penetration testing requirements
- Transparent incident response procedures
- Mechanisms for continuous stakeholder involvement (e.g., public consultations, oversight committees, periodic evaluations)

31. Specialised judicial generative AI models represent a critical requirement for effective AI deployment in legal contexts. Legal language constitutes a technical domain with specialised terminology that often diverges significantly from common usage. Terms such as *"consideration," "prejudice," "public order", or "bona fides"* carry specific legal meanings that differ substantially from their everyday definitions. This semantic complexity necessitates generative AI models trained specifically on legal corpora from each jurisdiction, ensuring accurate use of technical legal terminology and correct contextual interpretation.

32. Customised judicial generative AI solutions may incorporate:

- Integration with official legal databases and verified jurisprudential sources
- Jurisdiction-specific training on legal texts, court decisions, and procedural documents
- Multi-language capabilities reflecting the linguistic diversity of European judicial systems
- Specialised modules for different areas of law (civil, criminal, administrative, constitutional)
- Built-in citation verification systems cross-referencing official legal registries
- Automated detection and flagging of potential legal inconsistencies or anachronisms

[19] Council of Europe. (2018). Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data [Modernised Convention 108], Council of Europe Treaty Series - No. 223. Strasbourg.
[20] For example, Spain's State Technical Committee for the Electronic Judicial Administration (2024). Policy on the Use of Artificial Intelligence in the Administration of Justice. A courtesy translation to English can be found on: https://www.administraciondejusticia.gob.es/cteaje/normativa-complementaria.
[21] For example, Spain's Royal Decree 311/2022, which regulates the National Security Scheme.

33. Infrastructure sovereignty constitutes an essential consideration. Judicial generative AI systems should operate on infrastructure under the exclusive control of judicial authorities or governmental entities when opting for open-source models running on public servers, or in collaboration with supplier companies when opting for commercial models.

2.1.5. Ensure sound management in the deployment of generative AI systems

34. The implementation of generative AI systems in judicial administration should follow a comprehensive approach that includes several key elements: an incremental deployment strategy with appropriate change management processes and pilot phases with evaluations for review, careful consideration of licensing models (proprietary versus open-source technology), ongoing vendor support and training requirements, regular audits, monitoring, and review of AI system performance and associated risks.

35. To ensure better understanding of the technology, as well as to reduce potential barriers to adoption and facilitate change management, it is suggested that the implementation of generative AI in the administration of justice follows a gradual, step-by-step approach that is differentiated according to the complexity and nature of the issues being addressed:

Preparatory Phase: Before any implementation, institutional readiness audits should be conducted, basic staff training should be provided, safety protocols should be established, and oversight committees should be created.

First Phase: Limited to basic administrative and support functions that do not directly impact judicial decision-making. This phase should focus on document management and information processing, such as extracting relevant data from unstructured documents, obtaining key information (names, dates, amounts) and automatically incorporating them into predefined templates. Likewise, it could include basic workload management through analysis of historical patterns and assistance in case allocation. Basic information services and user assistance may also be implemented, such as procedural guidance systems in plain language for users and specialised chatbots for straightforward procedural enquiries. These should be complemented with a disclaimer on the risks involved. Non-simultaneous transcription of hearings could also be considered in this phase. Both courts and the administration of justice should evaluate workload impact, monitor user satisfaction, and measure adaptation time.

Second Phase: Expansion to more sophisticated analytical and drafting support functions. This phase should encompass advanced document management and information processing capabilities, including automatic generation of hearing minutes and linking arguments between different procedural writings. This phase could also extend to drafting support and judicial documentation, with automatic creation of standard procedural documents, assistance in drafting routine orders, and verification of document consistency. Likewise, with creation of jurisprudence repositories organised by case type and subject matter and automated recognition of disputed issues by comparing complaints and responses, as well as drawing up thematic and chronological indexes for voluminous case files. Case-law analysis and decision support should be introduced for preliminary analysis of appeals regarding compliance with formal and technical requirements. Courts should conduct A/B testing for efficiency comparison, evaluate decision bias, and analyse usage patterns.

Third Phase: Introduction of more advanced analytical and advisory functions, maintaining strict judicial oversight. This phase requires positive evaluation of previous phases, specific approval for each sub-phase, and clearly defined rollback protocols. This phase can be further subdivided into three sub-phases:

Initial Sub-Phase: Focus on case-law research and document organisation through analysis of complaints and responses with suggestions of related case law, and identification and systematically organised jurisprudence relevant to each case for judicial consideration. Drafting support should include generation of structured summaries of judgments, procedural documents, and appeals, requiring comprehensive judicial verification. Advanced case management functions could facilitate the identification of argumentative patterns in similar claims and detection of proceedings that may raise similar legal questions.

Intermediate Sub-Phase: Enhanced case-law research and document organisation to support judicial preparation, providing systematically organised legal precedents and relevant jurisprudence for judicial consideration. Drafting support could include preparation of preparatory documents such as reports, outlines, and mind-maps that present an overview of points of agreement and disagreement between parties.

Advanced Sub-Phase: In highly standardised procedures with well-established legal precedents, generative AI could assist by providing comprehensive case-law analysis with pattern recognition across similar cases, identifying procedural precedents and potential inconsistencies for judicial review, and proposing drafts that incorporate the relevant case law. Likewise, it could help draft structured templates for routine procedural documents based on established judicial practices, while mandatory human analysis and decision-making would remain in place for all legal and judicial determinations and the final approval of every output.

36. All of the above should be applied as pilot projects in payment order procedures, assessed claims for determined amounts, verbal trials of small amounts, standardised administrative fines, or labour claims for uncontroversial amounts, in which all evidence is documentary.

37. Matters affecting fundamental rights, involving minors or incapacitated persons, or entailing deprivation of liberty measures require full human judicial intervention, limiting generative AI to technical support functions.

2.1.6. Be aware of how the parties are using generative AI

38. Both courts and the administration of justice should establish clear protocols for monitoring and managing the use of generative AI by legal practitioners and parties in proceedings. This may include requiring parties and their legal representatives to disclose when generative AI systems have been used in the preparation of legal documents, case research, or evidence compilation, when deemed appropriate.

39. Procedural rules should specify the obligation to notify the use of generative AI in case preparation, with appropriate sanctions for non-disclosure or excessive reliance on AI-generated content without proper verification.

40. Courts should be particularly vigilant regarding multiple forms of AI-related misconduct, including: the submission of fabricated evidence, non-existent legal citations, or AI-generated content that has not been adequately reviewed by qualified professionals; attempts to manipulate AI systems through hidden prompts or instructions embedded within documents, such as invisible text, unusual formatting, or microscopic embedded instructions designed to

compromise AI-assisted analysis; and the use of generative AI to create fraudulent videos or graphic evidence. All such practices must be regulated both procedurally and criminally.

41. To maintain procedural fairness and equality of arms, courts should ensure that all parties have equal access to information about generative AI use in proceedings and that no party gains an unfair advantage through undisclosed AI assistance.

2.1.7. Ensure awareness and training of AI users

42. Judges, prosecutors, administrative staff, and all judicial personnel should receive adequate training to understand, manage, and critically evaluate the generative AI tools available to them and those on the market. Specialised training and heightened awareness of these tools' potential and limitations will ensure ethical and effective use of the technology, always preserving human judgment in decision-making.

43. The implementation of AI in the administration of justice should be characterised by transparency and automation. To ensure certainty and consistency, judges should not need to perform specific operations to activate the system; its use should be predetermined by procedural laws.

44. Judges' core competences should focus on maintaining and applying their critical judgment, a quality inherent in and expected of the judicial function, rigorously evaluating every outcome provided by AI. This entails avoiding a complacent or passive stance toward technological suggestions and instead adopting a critical, even dissenting, attitude toward certain alternatives or proposals generated by these systems.

45. If judges are to be equipped with generative AI systems with which they can interact directly, they should receive specialised training in their use (prompt engineering, context engineering, etc.). They should acquire a solid technical understanding of these models' inherent limitations: the scope of their capabilities, the areas in which they tend to fail, their strengths, and the types of queries that are not technically viable or appropriate.

46. Judges should be able to analyse and question generative AI outputs, identifying potential biases, inconsistencies, or limitations in the reasoning. They should develop criteria for assessing when and how it is appropriate to use generative AI in different types of cases and acquire the ability to move efficiently between traditional sources (such as case-law databases and doctrinal articles) and generative AI systems to obtain the most relevant and reliable information.

47. Judges should be trained to avoid the so-called "bubble effect," which arises when a judge, impressed by the apparent efficiency and accuracy of generative AI systems, develops excessive confidence in the case law, reports, or summaries they provide, risking abandonment of independent consultation of original case-law databases and disregard for the underlying documents.

48. A continuous training policy should be developed for new generative AI tools as they emerge, as well as for advances that AI itself makes in different fields affecting justice (image generation, videos, voices, etc.). Judicial institutions should establish systematic mechanisms to keep judicial professionals updated on emerging AI capabilities and their potential implications for the administration of justice. This should include understanding new forms of AI-generated evidence, detecting deepfakes and synthetic content, and adapting to evolving methods of digital manipulation that may affect judicial proceedings. Such training should be regularly updated to reflect the state of the art in generative AI technology and its applications

in legal contexts. To ensure consistency in application, standardised training curricula could be considered.

2.2. Normative aspects for generative AI in courts

49. The deployment of generative AI systems in judicial settings requires a thorough assessment against the fundamental legal principles established by European legal frameworks and national constitutional provisions. To safeguard compliance with these legal principles, the CEPEJ Assessment Tool[22] and the HUDERIA impact-analysis methodology[23] could be used.

2.2.1. Right to a fair trial and judicial independence (Article 6 ECHR)[24]

2.2.1.1. Fair trial guarantees

50. Generative AI systems may compromise procedural fairness through opaque decision-making processes, non-explainable automated recommendations, or unequal procedural advantages arising from differential access to advanced technological tools among parties.[25]

51. Technical safeguards derived from these provisions should include the implementation of explainable AI systems,[26] full traceability of AI reasoning processes,[27] equal access for all parties to AI-generated outputs, and technical mechanisms ensuring genuine equality of arms.[28] AI systems supporting judicial decisions should provide outputs that are both demonstrably explainable and reliable, and that are understandable to non-technical experts. They should also implement measures to minimise inaccuracies, ensure that correct legal reasoning is followed and protect data security.

2.2.1.2. Principle of legality, legal certainty and predictability

52. Generative AI systems that produce "hallucinations of law" or reference non-existent legal provisions pose serious risks to legal certainty and the rule of law.

53. The implementation of AI should not undermine the foreseeability and consistency of judicial decisions. The information sources, criteria, and methodologies used by generative AI models should be thoroughly documented and clearly justified to prevent the introduction of uncertainty or arbitrariness into judicial processes.

54. Technical safeguards should include automatic validation of legal citations, exclusive use of certified official legal sources, and alert systems for generated content that cannot be verified against authoritative legal databases.[29]

---

[22] European Commission for the Efficiency of Justice (CEPEJ). (2023, December 4). Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment. Council of Europe.

[23] Council of Europe, Committee on Artificial Intelligence. (2024, November 28). Methodology for the Risk and Impact Assessment of Artificial Intelligence Systems from the Point of View of Human Rights, Democracy and the Rule of Law (HUDERIA Methodology). Council of Europe.

[24] Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, ETS No. 5, 213 U.N.T.S. 221. https://www.echr.coe.int/documents/d/echr/convention_ENG

[25] Article 6 § 1 of the ECHR guarantees the right to a *"(…) fair and public hearing within a reasonable time by an independent and impartial tribunal established by law."*

[26] Level of understanding how the AI-based system came up with a given result, see ISO/IEC TR 29119-11:2020

[27] Requiring organisations to maintain detailed logs of AI operations, including data inputs, model decisions, errors and user interactions, see ISO/IEC 42001:2023

[28] The CEPEJ Assessment Tool identify as risk #8: *"Risk of unfair advantage for one party to the trial".*

[29] The CEPEJ Assessment Tool identify as risk #11: *"Risk of generation and use of inexistent legal provisions by generative AI".*

2.2.1.3. Exclusivity of judicial authority

55. Jurisdictional power lies exclusively with judges in accordance with constitutional and legal provisions. Decision-making, legal reasoning, and evidentiary assessment should not be delegated to generative AI systems; such systems may be employed solely for auxiliary or preparatory tasks that support, but never replace, the human judgment of the judicial body.

2.2.1.4. Mandatory human responsibility and supervision

56. Judges retain ultimate and exclusive responsibility for all judicial decisions, regardless of any technological assistance employed in the decision-making process. Judicial officers should fully understand the limitations of any generative AI system used and should thoroughly review all AI-generated outputs before incorporating them into judicial acts or decisions.

2.2.1.5. Judicial independence and non-binding nature of AI

57. Judicial independence[30] may be undermined when judges experience direct or indirect pressure to follow AI-generated recommendations, thereby constraining their freedom of judgment, legal interpretation, and autonomous decision-making processes. This independence should extend to freedom from technological coercion or systematic bias introduced through AI recommendations. Judges should retain the right to opt out of receiving AI suggestions.

58. Recommendations, suggestions, or outputs generated by AI systems shall never be binding on judges, nor shall their use be mandatory, thereby fully preserving judicial independence in legal interpretation, reasoning, and decision-making. Judges shall not be required to specifically justify their departure from proposals or recommendations generated by AI systems.

59. The design of generative AI models should not prioritise maximising their persuasive capabilities in order to minimise the risk of undue psychological influence and automation bias having a negative effect on judicial impartiality and freedom when evaluating AI-generated judicial opinions.[31]

60. Technical safeguards should be proportionate to system risks and include designing generative AI models strictly as non-binding auxiliary tools, eliminating any mechanisms that oblige or pressure judges to justify departing from algorithmic suggestions, performing rigorous evaluations prior to any deployed use of AI tools, conducting independent technical audits, and providing mechanisms and continuous training preventing automation bias to occur and to preserve autonomous human judgment.[32]

2.2.1.6. Transparency and adequate reasoning

61. It should be clearly documented if generative AI is used in any aspect of judicial decision-making.[33] Parties have the right to know which elements of a judicial resolution have been assisted by generative AI systems.

---

[30] Article 6 § 1 ECHR requires judgment by an *"independent and impartial tribunal established by law"*.

[31] Salvi, F., Horta Ribeiro, M., Gallotti, R., & West, R. (2024). On the Conversational Persuasiveness of Large Language Models: A Randomized Controlled Trial. arXiv preprint arXiv:2403.14380

[32] The CEPEJ Assessment Tool identify as risk #12: *"Risk of disempowerment and limitation of accountability of the judge through the use of non-explainable AI"*. HUDERIA's framework enables recurrent assessment of potential adverse impacts on judicial autonomy.

[33] European Commission for the Efficiency of Justice. (2025, January 31). Reflections of the AIAB on the use of artificial intelligence in judicial systems (CEPEJ-AIAB(2025)1Rev5).

2.2.1.7. Right of access to a court

62. This fundamental right faces potential compromise through automated systems that replace or hinder access to human judicial decision-making, affecting the right to human judgment in judicial matters.

63. Technical safeguards should guarantee meaningful access to a human judge, confine generative AI to strictly assistive functions, and prohibit total automation of judicial decisions.[34]

2.2.2. Right to liberty and security (Article 5 ECHR)

64. Decisions affecting personal liberty require the highest level of human oversight and justification. Generative AI systems that influence detention, pre-trial custody, or precautionary measures without adequate human review and transparent justification pose significant risks to fundamental freedoms.[35]

65. The introduction of AI in judicial functions should follow a progressive, incremental approach, initially limited to administrative tasks or standardised procedures with predictable outcomes. Generative AI accuracy, effectiveness, and safety should be rigorously evaluated before expanding to more sensitive judicial functions.

66. Technical safeguards should include generative AI systems that promote alternatives to imprisonment, mandatory human-readable explanations of AI-generated recommendations affecting liberty, and comprehensive human review of every AI-assisted decision where deprivation of liberty might be applied.

2.2.3. Right to privacy and data protection (Article 8 ECHR)

67. Privacy and the protection of personal data should be ensured if generative AI systems are used. All technology and data employed should remain under the exclusive control of the judiciary, so as to prevent unauthorised access, leaks, or external interference and to preserve the institution's sovereignty over the information. Judicial data should remain under the exclusive control of judicial authorities. Courts should maintain the capacity to change AI providers without loss of functionality as well as the right to verify algorithms and training data through ongoing audit processes.

68. Technical safeguards may include mandatory local pseudonymisation of personal data prior to loading this data into the AI system, execution of AI models on government-controlled infrastructure or in commercially available infrastructure; in the latter, mediated through formal contractual arrangements between public authorities and technology service providers, and implementation of strict data protection protocols.[36]

[34] The CEPEJ Assessment Tool identify as risk #6: *"Risk of AI replacing the access to the judge".*

[35] Article 5 § 1 ECHR protects the right to liberty subject only to narrowly defined, lawful exceptions. Article 5 § 2 establishes the right to be informed *"of the reasons for his arrest".* Article 5 § 4 guarantees the right to judicial review of detention's legality.

[36] The CEPEJ Assessment Tool identify as risk #2: *"Risk of personal data or trade secret disclosure"*. The HUDERIA privacy impact analysis framework specifically addresses risks of unauthorized disclosure of sensitive data through its COBRA risk analysis methodology.

2.2.4. Right to an effective remedy (Article 13 ECHR)

69. AI-powered automated systems should not hinder effective access to remedies or judicial review.[37]

70. Technical safeguards should ensure mandatory human review of AI-assisted decisions and establish specific remedial procedures for challenging automated or *"AI-influenced"* determinations.

2.2.5. Prohibition of discrimination (Article 14 ECHR and Protocol No. 12)

2.2.5.1. Non-discrimination principle

71. Algorithmic biases in generative AI systems pose significant risks of replicating or amplifying existing discrimination across protected characteristics, potentially violating core equality principles.[38]

72. Technical safeguards should include periodic bias audits, data cleaning and balancing, fairness regularization terms during training (such as demographic parity or equalized odds constraints), counterfactual/A-B testing, and continuous monitoring of discrimination metrics.[39]

2.2.5.2. Neutrality and active bias prevention

73. AI systems should be periodically evaluated to detect and mitigate bias, ensuring they do not perpetuate existing discriminations or create new forms of inequality in judicial decision-making. Systematic monitoring for bias across all protected characteristics is mandatory.

2.2.6. Additional Safeguards

2.2.6.1. Explainability and transparent reasoning

74. AI systems used in the judicial sphere should provide clear and understandable explanations of the path followed for each result or recommendation, so that judicial operators can critically verify them. To this end, the systems should disclose their internal inference process in a transparent and auditable manner, ensuring substantive alignment and avoiding superficial compliance simulations that mask their real logic.[40] Systems should be able to indicate the specific sources of each generated statement. Only interpretable models that reveal their internal decision-making processes should be employed.

2.2.6.2. Proportionality and caution in generative AI deployment

75. The introduction of AI in judicial functions should follow a progressive, incremental approach, initially limited to administrative tasks or standardised procedures with predictable

[37] Article 13 ECHR establishes that *"Everyone whose rights and freedoms as set forth in this Convention are violated shall have an effective remedy before a national authority notwithstanding that the violation has been committed by persons acting in an official capacity."*.

[38] Article 14 ECHR establishes that "The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status."

[39] The CEPEJ Assessment Tool identify as risk #10: *"Risk of discrimination or amplification of discrimination"*. HUDERIA's stakeholder engagement process enables identification of differentiated impacts on vulnerable groups.

[40] Greenblatt, R., Denison, C., Wright, B., Roger, F., MacDiarmid, M., Marks, S., Treutlein, J., Belonax, T., Chen, J., Duvenaud, D., Khan, A., Michael, J., Mindermann, S., Perez, E., Petrini, L., Uesato, J., Kaplan, J., Shlegeris, B., Bowman, S. R., & Hubinger, E. (2024). Alignment faking in large language models [arXiv preprint arXiv:2412.14093v2]. https://doi.org/10.48550/arXiv.2412.14093

outcomes. Generative AI accuracy, effectiveness, and safety should be rigorously evaluated before expanding to more sensitive judicial functions.

2.2.6.3. State liability for damages caused by generative AI use

76. The State should assume responsibility for any harm caused by judicial errors or failures of generative AI systems, in line with existing principles of governmental liability. Neither the autonomy nor the technological complexity of generative AI should be invoked to exempt the State from its duty to provide redress.

2.2.6.4. Principle of procedural fairness and efficiency

77. The introduction of AI into the administration of justice should not result in a deterioration of procedural safeguards or an impairment of effective judicial protection; on the contrary, it should strengthen procedural rights and judicial effectiveness. For example, it should improve access to a court, offer better access to evidence or improve the mean's to better present one's case. Furthermore, generative AI implementation should result in measurable improvements in response time, communication clarity, decision consistency, and case management efficiency. Courts should maintain the capacity to return to previous systems if generative AI does not objectively improve judicial services. Regular evaluations should assess compliance with this principle.[41]

3. GLOSSARY

Principle of technological neutrality

Is generally described as the freedom of individuals and organizations to choose the most appropriate and suitable technology to their needs and requirements for development, acquisition, use or commercialisation, without dependencies on knowledge involved as information or data.[42]

Fairness regularizers during model training

Specific mathematical techniques such as:

- Demographic parity regularizers: penalize the model when prediction rates differ significantly across demographic groups
- Equalized odds regularizers: ensure similar true positive and false positive rates across protected groups
- Individual fairness constraints: guarantee that similar individuals receive similar treatment
- Adversarial debiasing: use adversarial networks to remove information about protected characteristics from learned representations

[41] Evaluation indicators should include user satisfaction surveys, performance metrics comparison with pre-AI systems, stakeholder feedback from judges, court staff, lawyers and citizens, and assessment of procedural timeline improvements.

[42] See: https://icannwiki.org/Technology_neutrality