



Centro de Estudios de Justicia de las Américas
Justice Studies Center of the Americas

Artificial Intelligence in the Judiciary: Reflections and Guidelines for the Americas

www.cejamericas.org



© 2026 Justice Studies Center of the Americas (JSCA/CEJA)

San Pío X 2460, Office 608, Providencia, Santiago, Chile

Phone: +(56 9) 5611 2162

www.cejamericas.org

Title: *Artificial Intelligence in the Judiciary: Reflections and Guidelines for the Americas*

First Edition, 2026

Copyright and ISBN: Pending.

Author

Pablo Pizarro Zúñiga

Project Team

Nataly Ponce Chauca

Rodrigo Nakada Castro

Kevin Gutierrez Parra

This publication is distributed free of charge in both print and digital formats.

The reproduction of this publication, in whole or in part, is authorized provided that the source is properly acknowledged

Introduction	3
I. Artificial Intelligence in the Judiciary	5
The Integration of Artificial Intelligence into the Administration of Justice	5
II. Guidelines for Integrating Artificial Intelligence into the Judiciary	16
Problem Identification	16
Expected outcomes	36
Conclusions	37
References	39

Introduction

In a rapidly transforming digital landscape fueled by artificial intelligence (AI), judicial authorities and other justice sector institutions are swiftly integrating these technologies into their management models and daily operations. Such innovations are increasingly explored as a means to enhance productivity ([Ribeiro & Carneiro 2025](#)) and facilitate administrative workflows. This trend reflects both the inherent potential of AI and the high expectations placed upon it. These systems have diverse applications within the justice system, including user communication, process optimization, data management, the systematization of jurisprudence, support for judicial functions, and even the automation of decision-making.

This breadth reflects not only technological versatility but also varying levels of implementation risk. These range from low-risk administrative uses that offer significant improvements to prohibited or high-risk applications that may jeopardize judicial functions, the fundamental rights of litigants, and, ultimately, the institutional legitimacy of the judiciary.

In their most intensive forms, AI systems risk distorting or entirely replacing judicial discretion. Even in more limited support roles, substantive dimensions of the legal process are affected, including the interpretation of facts, legal reasoning, the drafting of judgments, and the ability of parties to effectively understand and challenge the grounds of a decision. Additionally, uncertainty remains regarding their impact on the development of case law, legal practice, and the agency of judges ([CEPEJ 2018](#); [Münch & Ferraz 2024](#)).

Debates concerning the risks AI poses to fundamental rights—such as data privacy, access to justice, equality before the law, and the right to a human judge—frequently occur in the absence of comprehensive regulatory frameworks. Furthermore, most judiciaries in the Americas still lack governance structures and clear guidelines defining the principles and limits that should govern such implementation.

In this rapidly evolving scenario, this report contributes from two perspectives. The first chapter analyzes the potential and risks of AI in both administrative management and judicial decision-

making. The second chapter provides practical guidelines for introducing AI into the judiciaries of the Americas, based on these reflections. These guidelines adopt a public policy analysis and design approach ([Ivãlua 2020](#)), emphasizing the necessity of establishing robust governance mechanisms for these technologies.



Artificial Intelligence in the Judiciary

The Integration of Artificial Intelligence into the Administration of Justice

The expansion of artificial intelligence (AI)¹ marks an unprecedented shift in the delegation of complex functions to automated systems². These technologies process vast volumes of data with high speed and are capable of iterative learning, adapting to new variables, generating recommendations, and executing decisions based on algorithmic patterns ([Castets-Renard 2021](#)). These capabilities have accelerated their adoption across an increasing number of jurisdictions, within both managerial tasks and judicial functions³.

¹For the purposes of this report, "artificial intelligence systems or tools" are defined as machine-based technologies that process data using algorithms and computational models to generate results such as recommendations, predictions, decisions, or content. These systems operate with varying degrees of autonomy and can be adapted after implementation. They are designed to emulate certain human cognitive functions, such as analysis, inference, and decision-making. However, this does not imply an equivalent reproduction or understanding of content as occurs in human reasoning.

²Automated decision-making systems are tools designed to replace, either partially or fully, decision-making processes that would otherwise be carried out by humans. These systems operate using automated reasoning techniques that do not necessarily involve AI. They analyze large volumes of data to identify patterns, correlations, and inferences that are useful for decision-making ([Castets-Renard 2021](#); [UNESCO 2023](#)).

³ As noted in the chapter on the administration of justice in the [OECD document Governing with Artificial Intelligence](#), and as will be seen throughout this document, there are multiple examples of AI applications in

Given its rapid evolution in task performance, accuracy, and efficiency—coupled with the exponential growth projected for the coming years—characterizing AI as a "technological disruption" is fully justified. AI applications streamline repetitive tasks and facilitate information management, thereby reducing case processing times. Furthermore, empirical research attributes additional benefits to these systems, such as mitigating discrimination and bias in substantive decisions and promoting greater judicial consistency ([Kleinberg et al. 2018](#)).

Uses for the administrative management of judicial services

AI represents a strategic opportunity to deepen the institutional modernization of judiciaries in the Americas, particularly regarding the administrative management of judicial services. Implementing AI in management helps distinguish administrative functions from purely adjudicative ones. Automating administrative support tasks fosters the development of professional profiles specialized in management—requiring expertise in computer science, data analysis, and technology oversight—which are distinct from traditional legal training. This technical professionalization allows judges to decouple from administrative management and focus exclusively on their primary responsibility of dispute resolution.

Additionally, reallocating institutional resources away from routine tasks enables the reassignment of personnel to higher-value functions, substantially enhancing service quality through personalized user attention and in-depth legal analysis. Moreover, the adoption of these technologies has the potential to address systemic deficiencies, such as the lack of data interoperability between institutions.

the judicial field worldwide. For example, Argentina uses Prometea to automate judicial tasks, and Austria and Croatia have developed automated decision anonymization systems. Brazil has VICTOR AI, Chat-JT, MARIA, and RAFA 2030; Bulgaria, Greece, and Romania have the i-ACCESS My Rights chatbot, which provides information to children and adolescents; and Colombia has PretorIA. Colombia implemented Pretor IA; Slovenia uses AI for automated transcriptions; and Spain works with tools such as VioGén and Riscanvi, which estimate recidivism probability similarly to COMPAS in the United States. Peru uses Amauta Pro to draft resolutions on protective measures in cases of violence.

From an access-to-justice perspective, the versatility of these technologies enables the rapid and cost-effective piloting and implementation of new solutions, effectively reducing barriers through remote services. This is particularly significant in contexts where geographical constraints impose disproportionate costs on court access.

Administrative management functions are classified as low-risk under the risk-based framework of the [2024 European Union AI Act](#). Consequently, systems intended for document management, hearing scheduling, or the anonymization of judicial decisions are subject to less stringent regulatory obligations. From the standpoint of balancing fundamental guarantees with institutional efficiency, the implementation of these systems is highly recommended. They represent ideal instruments for addressing systemic inefficiencies—such as case backlogs, procedural delays, and accessibility barriers—through the integration of novel technological resources.

Uses in Adjudication

The context shifts significantly when AI is applied to adjudicative functions. Currently, no clear consensus exists regarding the risk levels associated with AI in this domain. This ambiguity stems from several factors, including a regulatory vacuum in the judicial sphere, the absence of institutional guidelines in many jurisdictions, and varying risk assessment methodologies across different judiciaries. Furthermore, identifying the threshold where technical assistance constitutes an undue encroachment upon the exclusive functions of the judge—or replaces human decision-making entirely—remains a complex challenge. Consequently, there is a high degree of discretion in the adoption of these technologies ([Buchholtz 2019](#)).

At present, empirical evidence concerning the impact of AI on judicial outcomes, procedural guarantees, privacy, and the development of jurisprudence remains scarce and inconclusive. This lack of certainty complicates evidence-based policy-making regarding the regulation or prohibition of specific applications ([Dias & Satiro 2024](#)). Therefore, prioritizing solutions that safeguard fundamental rights is a necessary strategy for guiding the adoption of these technologies ([Perez-Pacheco & Simón Castellano 2025](#)).

Despite these uncertainties, it is evident that incorporating AI into adjudication carries structural risks. Algorithms are not neutral or objective; they are human constructs that reproduce and may

even amplify the biases, errors, and discriminatory patterns inherent in their training data ([Chiao 2019](#); [Cofone & W Khern-am-nuai 2024](#); [Ferrara 2023](#); [Malek 2022](#); [West 2020](#)).

This risk underscores the urgent need for robust institutional oversight, as courts possess the discretion to authorize, design, or procure AI systems that substantially influence decision-making. It is imperative for the judiciary to establish governance frameworks and clear rules for all stakeholders. Such measures are essential to prevent applications incompatible with due process guarantees, the judicial duty to adjudicate, and the prohibition against delegating adjudicative functions. Currently, the absence of oversight allows commercial AI tools to operate without restriction in highly sensitive areas, even though these same applications are prohibited or strictly regulated in other jurisdictions due to their high-risk nature ([Vargas-Murillo et al 2024](#)). This creates significant disparities in regulatory permissiveness across regions.

In contexts characterized by limited resources, weak governance structures, and systemic judicial congestion, the incentives to adopt AI without adequate safeguards are heightened ([JSCA 2025](#)). The pressure to mitigate workloads may lead to the acceptance of risks that would be considered unacceptable in other environments. To facilitate risk analysis, two levels of adjudicative tools are distinguished: (1) judicial decision support, and (2) the automated generation of content or the substitution of judicial discretion.

Decision Support Systems in Adjudication

These AI applications are designed to assist with critical functions, such as the interpretation of legal rules and the evaluation of evidence. Although frequently classified as low risk, these technologies do more than merely accelerate legal workflows; they fundamentally alter how relevant information is prioritized and processed for decision-making. By providing algorithmic summaries and outcome predictions, these systems have the capacity to influence judicial deliberation. Consequently, the improvements in efficiency and accuracy must be weighed against the risk of introducing biases stemming from training data or design parameters, which could ultimately compromise judicial impartiality.

One application of particular interest within the judicial sphere is automated document summarization. These tools are compelling because documentary evidence often comprises voluminous texts—including judgments, contracts, expert reports, transcripts, and regulations—the review of which consumes significant institutional resources. An automated summarization tool can reduce tasks that formerly required weeks to mere seconds, substantially mitigating workloads. In this regard, integrating such technology could provide a significant contribution to the daily operations of the justice system.

However, summarization is not merely a process of abridgment; it necessitates determining which textual fragments are relevant—a task that depends heavily on the specificities of each case. For instance, the relevance of information varies according to the subject matter, the nature of the alleged offense, or the characteristics of the parties involved. Therefore, if such a tool is not rigorously designed, it may omit essential information or reproduce biases during the selection of "relevant" content.

Given these challenges, transparency and effective accountability mechanisms are essential for establishing AI systems aligned with human rights standards that can effectively assist judges in decision-making ([Coglianese & Lehr 2018](#); [McGregor et al 2019](#)). Although these support tools are governed by the principles of non-substitution of human discretion and human oversight in jurisdictions where guidelines exist, such supervision may be impractical. Negligent review of outputs or excessive workloads often prevent effective oversight, leading to "automation bias." This raises significant questions regarding the legitimacy of decisions influenced by algorithmic codes that are not subject to democratic control, suggesting a need for rigorous analysis and evaluation of these instruments. Both the [European Union's AI Act](#) and the [Guidelines of the Superior Council of the Judicature of Colombia](#) adopt this approach, classifying applications that support critical judicial functions—such as the interpretation of facts and norms, the evaluation of evidence, criminal profiling, and the prediction of recidivism or victimization risks—as high-risk, even when human intervention is involved.

Generation of Judicial Resolutions⁴ and the Substitution of Discretion

The second level encompasses applications that substitute human discretion with automated decisions, generate novel content for judgments, or produce unsupervised legal outputs⁵. Although most guidelines and expert opinions examined for this report oppose the replacement of judges by AI, some perspectives identify potential in substitution scenarios⁶. In practice, despite formal programmatic statements prohibiting substitution or affirming the primacy of human will, several judiciaries employ these technologies to automatically supplant judicial discretion in repetitive, low-complexity cases, interlocutory rulings, and admissibility assessments⁷.

⁴ This category excludes generative AI tools utilized for drafting refinements, orthographic correction, or the implementation of plain language. Instead, it specifically refers to outputs involving statutory interpretation, evidentiary assessment, decision proposals, and any application understood to substitute tasks intrinsic to the essence of the adjudicative function.

⁵Experiences within the region diverge regarding the use of generative AI for the issuance of resolutions. For instance, the [Agreement of the Superior Council of the Judicature of Colombia](#) prohibits its use for the assessment of evidence, factual scrutiny, the formulation of value judgments, and the resolution of legal issues. Furthermore, it bars the application of norms, the provision of legal reasoning, or the adoption of decisions based exclusively on the tool's outputs. Conversely, the "[Pilot Program for the Strategic and Responsible Use of Generative AI in the Argentine Justice System](#)" tested and validated a broader range of applications, including: the interpretation and application of law (statutes and principles); document drafting (judgments, resolutions, and opinions); evidentiary assessment (testimony analysis and substantiation); and information retrieval and analysis (case law, legal doctrine, and the evaluation of rulings).

⁶ According to [Alarie \(2016\)](#), AI holds the potential to eliminate the "gray areas" of the law—that is, the interpretative realm subject to human discretion. In this regard, [Alarie et al. \(2018\)](#) argue that AI predictions would be more objective, rapid, and precise, offering the potential to reduce the margin of error and the cognitive biases inherent in human analysis.

⁷For a detailed analysis of specific applications, it is recommended to consult the report "[Artificial Intelligence Applications in the Judiciaries of Ibero-America](#)" by the Ibero-American Judicial Summit (2023). This document catalogs 97 AI solutions implemented across the region, with the majority located in Brazil (57) and Spain (23), followed by Peru, Puerto Rico, Chile, the Dominican Republic, Uruguay, Costa Rica, and Colombia. The report classifies these solutions into various categories, with Machine Learning being the most widely implemented type, followed by Deep Learning tools. The document serves as a valuable repository for observing the practical uses assigned to these technologies, which include: the automation of legal and administrative processes; virtual assistance for legal professionals and citizens; enhancements in forensic investigation and evidentiary analysis; the optimization of case management and resource allocation; anonymization and classification; gender-based violence risk prediction; complaint classification; and the estimation of settlement probabilities, among others.

One could argue that such substitutive uses are incompatible with a judicial function that respects the democratic rule of law, regardless of the procedural stage or case type. The core responsibilities of human judges remain irreplaceable ([Gutiérrez 2024](#)). Current AI systems lack the capacity to critically evaluate the veracity of facts, exercise normative reasoning, demonstrate empathy, or innovate—skills that are indispensable for issuing judgments and conducting hearings.

Although these tasks are often delegated to clerks or legal assistants in practice, procedural rules explicitly vest the power of decision—such as determining admissibility—in the judge, as it is fundamental to the realization of the right of access to justice. Any modification, whether human or technological, that divests the judge of powers expressly granted by law must be addressed through transparent democratic legislative debate.

Furthermore, delegating decision-making to these systems risks compromising the evolution of jurisprudence ([CEPEJ 2018](#)). Because they are trained on historical data, these models lack the capacity to apply rules to unprecedented situations, balance fundamental rights during states of emergency, or evaluate institutional checks and balances amidst novel political crises.

Entrusting adjudication to these systems can cause jurisprudential stagnation or regression, creating a recursive feedback loop. Currently, algorithms are predominantly trained on human decisions. To update these models to reflect contemporary social phenomena and current regulatory frameworks, they require recent training data. However, if the judicial system relies on automated adjudication, future training will be based on machine-generated outputs, thereby entrenching the historical criteria upon which the original models were trained. This precludes the dynamic interaction between norms and principles that underpins judicial decision-making and case-by-case interpretation.

Compounding this issue is a structural risk to the counter-majoritarian role of the courts. Algorithmic outputs tend to converge on the statistically most probable outcome, a phenomenon known as "majority bias." This logic is fundamentally incompatible with the mandate to protect the rights of minorities and vulnerable groups—precisely those who require judicial intervention to guarantee their fundamental rights.

Critical Debates Regarding AI Integration in the Justice System

Models of Technological Provision and the Scope of Authorized Tools

In planning the integration of AI into the justice system, critical determination concerns the model of technological provision. This entails choosing between open-source models, in-house development, bespoke institutional tools, or commercial chatbots.

This decision extends beyond budgetary capacity or technical specifications, as the selected model holds substantive implications for judicial independence, information security, and transparency. The sensitivity of judicial data and its utilization in algorithmic systems mandate heightened precautions. Best practices, such as ensuring data anonymization prior to training and regulating data protection during the use of generative AI, constitute minimum standards for security and protection.

Furthermore, when prompts or inputs are managed by external providers, there is a distinct risk that data will migrate outside institutional infrastructure to be hosted on external servers. The removal of data from the sphere of institutional custody, combined with the conditions of access and exploitation by third parties, may prove incompatible with prevailing data protection regulations, the judiciary's information security policies, or statutory obligations regarding the secrecy and confidentiality of judicial files and ongoing criminal investigations. Consequently, a rigorous analysis of these legal instruments is required prior to final decision-making.

Moreover, the chosen architecture impacts the traceability, explainability, and publicity of decisions, as well as the mechanisms for operational verification and error correction. In-house development—whether autonomous or in collaboration with third parties—enables rigorous control over training data, facilitates bias mitigation, and ensures the compatibility of these tools with current regulations and fundamental rights.

Approaches within Latin America diverge significantly on this issue. In Colombia, the Superior Council of the Judicature has expressly prohibited the use of generative AI tools not officially authorized by the judiciary, thereby excluding the direct use of commercial solutions ([CSJ Colombia 2024](#)). In Brazil, the judiciary permits third-party solutions only if specifically tailored to

court needs and compliant with strict transparency, security, and institutional control requirements (CNJ Brazil 2025), though in practice, in-house developments are preferred (Ibero-American Judicial Summit, 2023). Conversely, in the Province of Chubut, Argentina, the agreement issuing guidelines on the ethical use of generative AI neither specifies the technical architecture nor references a system procured by the provincial judiciary, implying the tacit authorization of commercial chatbots (PJ Chubut 2025).

Impacts of AI on the Workplace

The integration of AI into the courts is not merely a technological challenge; it is a process of organizational transformation that directly impacts the workforce. From a change management perspective, it is essential to anticipate and responsibly address the shifts these tools introduce into work dynamics, ensuring they are not perceived as threats to job security or professional autonomy.

Previous research warns of the consequences of neglecting this human dimension. Bayne & Christin (2020) Identify strong internal resistance to the implementation of crime prediction tools. These technologies were rejected by court personnel not only because they were viewed as substitutes for traditional human functions but also because they were interpreted as mechanisms to intensify managerial control and surveillance. The study concluded that this workplace unrest ultimately compromised the quality of data collection, limited the effective use of algorithmic recommendations, and altered the exercise of discretion by decision-makers.

To mitigate these risks, the adoption of AI must follow an inclusive approach that acknowledges and respects the varying levels of digital literacy within the judicial system. The digital divide must not become a driver of professional exclusion or discrimination. Therefore, training programs featuring tailored strategies, personalized mentorship, and accessible formats should be implemented to ensure that all personnel, without distinction, possess the necessary conditions to adapt to this transition.

Conclusion: Towards a Strategic and Differentiated Implementation

The integration of AI into the justice system requires a strict distinction between administrative management, decision support, and the substitution of adjudicative functions. While technology offers a secure, low-risk avenue to address backlogs and improve access in administrative tasks, its application in adjudicative functions carries structural dangers that, if not properly regulated, may compromise due process. Operational efficiency cannot be achieved at the expense of fundamental guarantees or the human dimension of justice.

Consequently, the incorporation of AI demands a structured adoption process in which judicial governance bodies assume a leading role, preventing the discretionary and unsupervised deployment of high-risk applications. The absence of guidelines risks widening existing gaps in access to justice, exacerbating systemic bottlenecks, and eroding public confidence. Furthermore, without clear direction, spontaneous and unsupervised uses are likely to proliferate. In this context, institutional intervention emerges as a strategic imperative aimed at strengthening system legitimacy and mitigating the risks associated with a fragmented organizational response. Establishing a clear framework offers certainty to operators, litigants, and citizens, promoting an AI ecosystem that is transparent and consistent with the judiciary's objectives and obligations.

Accordingly, Chapter II provides guidelines for a thoughtful, responsible AI adoption process. These guidelines focus on high-impact interventions that introduce substantive improvements to judicial activities while preserving institutional legitimacy and minimizing potential impacts on the fundamental rights of those accessing the justice system.

II

Guidelines for Integrating Artificial Intelligence into the Judiciary

Problem Identification

Guiding Questions for Defining the Scope of Intervention

The initial step in planning the integration of AI into judicial activity involves clearly defining the intervention's purpose—specifically, the rationale for implementation and the problems it aims to resolve. To achieve this, the judicial organization must assess its current state and project the desired future state following AI regulation and implementation. This process entails addressing a series of strategic questions to guide the judiciary's response.

The following inquiries serve as a starting point for this discussion:

- What are the primary objectives of incorporating AI?
- Do excessive processing delays constitute a critical issue?
- Are channels for citizen access and communication with the courts insufficient?
- Are enhanced measures required to protect the privacy of the parties?
- Is information security management a current vulnerability?
- Is there a need to improve the accessibility and utilization of judicial data?

- Does the unsupervised or unauthorized use of generative AI pose a risk?
- Is AI a legitimate and proportionate means to address inconsistencies in judicial decisions?
- Do we currently measure judicial performance effectively and efficiently?

These inquiries demonstrate that any regulation or implementation strategy must be inextricably linked to a specific problem or unmet need. Determining *what* needs to be resolved will drive decisions regarding:

- The specific tools to be deployed;
- The levels of justice administration involved;
- The types of activities targeted;
- The budgetary allocation;
- The intended user base; and
- The beneficiaries of the implementation.

Accordingly, organizational reflection must delve into the root causes of the identified problems. This diagnosis should be grounded in the analysis of case processing data, performance indicators, comparative assessments of judicial activity, internal consultations, interviews, and focus groups, among other data collection mechanisms. This rigorous approach ensures the identification of root causes and the assessment of AI's suitability as a solution.

Leveraging this insight, the judiciary will be better positioned to identify comparative experiences that offer valuable lessons: which AI tools are most suitable, proportionate, and cost-effective for addressing identified needs; which are appropriate for solving distinct problems; and which tools utilized in other jurisdictions are incompatible with the applicable international, constitutional, or legal frameworks of the local context, rendering their replication inadvisable.

Establishing objectives

Establishing general and specific objectives is essential because it provides direction, consistency, and evaluation criteria. Having well-defined objectives allows for the alignment of different institutional actors, facilitates the allocation of resources, establishes clear priorities, and designs indicators to evaluate progress. In addition, specific objectives help to operationalize the policy by identifying concrete areas of intervention and realistic deadlines.

Based on a review of guidelines, recommendations, and policy instruments specific to the field of justice, a list of general and specific objectives was constructed that best reflects the proposals and statements contained therein, with the aim of guiding future discussions within the judicial branches and the adoption of their own objectives regarding the implementation and governance of AI⁸. The list consists of seven general objectives and twenty-five specific objectives. The latter are organized into four thematic dimensions: Operational and Efficiency, Support for the Judicial Function and Quality Improvement, Access and Transparency, and Governance and Implementation.

General Objectives

- Ensure the ethical, responsible, and secure use of AI.
- Guarantee the respect, protection, and promotion of fundamental rights.
- Prevent the displacement of human rationality by AI, thereby preserving critical judgment, legal reasoning, and substantive decision-making by individuals.
- Enhance the efficiency and quality of the administration of justice.
- Promote equity, objectivity, and non-discrimination by mitigating the influence of bias.
- Improve access to justice and communication with system users.
- Bolster transparency and public trust in the judicial system.

⁸The systematized objectives are derived from those observed in a comprehensive set of national and international documents. These include guidelines issued by the Judiciary of the Province of Chubut, the National Council of Justice of Brazil, the Judiciary of the Autonomous City of Buenos Aires, the General Secretariat of the CTEAJE (Spain), the Judiciary of New Zealand, the UK Judiciary, and the Judicial Councils of Canada and Colombia. Consideration was also given to UNESCO's draft guidelines, the 2025 Report of the UN Special Rapporteur on the independence of judges and lawyers, the CEPEJ Ethical Charter, and UNESCO's Global Toolkit on AI and the Rule of Law for the Judiciary.

Specific Objectives

Operational Efficiency and Management

- Accelerate judicial proceedings and reduce processing times through technological support.
- Automate repetitive, routine, or ancillary tasks to liberate the operational capacity of judicial personnel.
- Optimize case management and resource allocation to improve administrative efficiency.
- Support strategic planning through data processing and automated analysis.
- Efficiently process and analyze large volumes of legal and administrative data.
- Facilitate the access, retrieval, and utilization of relevant data during judicial proceedings.
- Automatically classify documents, files, lawsuits, and appeals based on similarity criteria or other variables.
- Assist in jurisprudential, evidentiary, or regulatory analysis, supporting interpretive work without supplanting it.
- Mitigate human error in document and procedural management through automated validation and verification tools.

Support for the Judicial Function and Quality Enhancement

- Assist judges in the analysis and synthesis of relevant information to facilitate decision-making.
- Strengthen the system of judicial precedents, promoting consistency in the resolution of analogous cases.
- Improve the clarity and quality of judicial drafting.
- Provide probability estimates for the evaluation of alternative scenarios (e.g., recidivism risk, reconciliation probability).

Access and Transparency

- Provide basic legal guidance and assistance to citizens and legal professionals, improving the user experience.
- Streamline the filing of lawsuits and procedural requirements.
- Protect user privacy through mechanisms for the anonymization of personal and sensitive data.
- Generate legal documents in plain language adapted to diverse audiences to strengthen transparency.
- Enable both internal users and external actors to interpret AI outputs, facilitating the review or challenge of these decisions.

Governance and Implementation

- Promote continuous training for judges and judicial staff regarding the informed and responsible use of AI.
- Establish permanent audit and monitoring mechanisms.
- Verify compliance of AI applications with prevailing regulatory frameworks.
- Ensure the explainability, transparency, and traceability of results generated by AI systems.
- Foster inter-institutional collaboration and the development of interoperable solutions.
- Promote the exchange of best practices to avoid the duplication of technological efforts.
- Ensure data protection, information security, and confidentiality, particularly regarding personal and sensitive data.

Available resources

Potential Applications of Artificial Intelligence in the Administration of Justice

Following the identification of problems and the definition of objectives, resource analysis requires a detailed mapping of the AI applications that best address the detected needs. Based on the document review, various AI applications in the judicial sphere are categorized into four key domains: Administrative Management, Judicial Support and Legal Analysis, Access to Justice, and Data Protection.

Administrative Management

- Automated classification of files and documents.
- Algorithmic case allocation based on the anticipated complexity of the matter and personnel availability.
- Management of electronic notifications.
- Assistance in the management and analysis of data derived from case processing systems.
- Automation tools for routine administrative tasks.
- Optimization of document and procedural management workflows.
- Digitization of physical files and documents.
- Automated information extraction for report generation.
- Scheduling and calendar management.
- Automated detection of omissions in documentation.
- Transcription of hearings.
- Image-to-text conversion (OCR).

Judicial Support and Legal Analysis

- Analysis and summarization of information relevant to decision-making.
- Information extraction from files, reports, and case law for the automated generation of legal memoranda.

- Identification and analysis of legal precedents.
- Enhancement of legal research and information retrieval systems.
- Pattern recognition tools to detect and group similar filed or pending cases.
- Forensic authentication of digital evidence (e.g., deepfake detection).
- Aggregate analysis of past decisions to identify jurisprudential trends.
- Assistance in drafting resolutions and structuring legal arguments.
- Verification of compliance with admissibility requirements.
- specialized case law search engines.
- Simulation of conciliation/settlement scenarios.
- Prediction of procedural outcomes.
- Risk assessment regarding recidivism.
- Automated judicial decision-making systems.
- Construction of quantum scales for calculating damages and compensation.

Access to Justice

- Conversational agents (chatbots) providing basic legal guidance.
- Service-oriented chatbots offering general information (hours, location, communication channels).
- Automated handling of inquiries (e.g., issuance of copies, certifications).
- Smart forms for filing lawsuits.
- Plain language translation tools for legal texts.
- Automated multilingual translation.
- Universal accessibility tools (text-to-speech, speech-to-text).
- Digital document validation.

Data Protection

- Anonymization applications for court rulings and algorithm training datasets.
- Identity protection tools for victims during trials or when presenting audiovisual evidence.
- Security enhancement applications for case management platforms.
- Algorithmic Impact Assessments (AIA).

Models of Technological Provision

Once the specific applications and objectives of AI have been defined, the mechanisms for its acquisition and deployment must be determined. It is recommended to evaluate the provision strategy from a multidimensional perspective, considering three fundamental axes: the available development models, cost analysis, and the associated operational and ethical implications.

Model Options

- In the process of incorporating AI, at least four distinct technology provision models can be identified:
- Open-Source Solutions: Typically developed by third parties, these tools offer the flexibility to be adapted locally to meet specific institutional needs.
- In-House Development: Projects executed by the judiciary's own technical teams, affording the institution greater control over design, implementation, and data governance.
- Bespoke External Development: Solutions custom-designed by external providers to meet the specific, unique requirements of the institution.
- General-Purpose Commercial Systems: Off-the-shelf tools (e.g., Gemini, Claude, ChatGPT) deployed with minimal customization.

Implementation Costs

The primary implication of this strategic decision is economics. While the adoption of open-source tools typically involves lower upfront costs, it demands significant internal technical capacity for adaptation, maintenance, and ongoing support. Conversely, bespoke developments commissioned from third parties generally incur higher costs due to the customization required and the associated support and update services. Understanding these cost differentials is not only relevant for comparative analysis but is a prerequisite for effective budget planning. Accurate cost projections enable informed decision-making, the appropriate allocation of budgetary resources, and the evaluation of the long-term viability of each model ([Ibero-American Judicial Summit 2023](#)).

Furthermore, this data facilitates the assessment of the cost-benefit ratio across available technological alternatives⁹.

Operational Implications of Provision Models

- Judicial authorities are advised to rigorously evaluate the benefits, risks, and limitations of each alternative, specifically considering the following elements:
- Data security and sovereignty.
- Traceability of the training lineage.
- Reliability and accuracy of outputs.
- Mitigation of hallucination risks.
- Capability for fine-tuning to enhance effectiveness.
- Compliance with regulatory frameworks.

Complementary Resource Requirements

The implementation process requires the allocation of multidisciplinary teams possessing technical, legal, and ethical expertise to oversee the technology lifecycle (or, alternatively, external technical consultancy). Essential resources include assessments of the existing technological infrastructure, capacity-building programs, and robust evaluation and monitoring mechanisms—such as Algorithmic Impact Assessments (AIA), external audits, bias analysis, and public consultations. Finally, a dedicated budget must be secured to establish internal governance structures and for the recruitment and creation of specialized professional profiles. These requirements will be explored in greater depth in the subsequent section regarding the implementation of roadmap.

⁹ A significant current limitation is the scarcity of comparable data regarding the actual investments made by other judiciaries in the deployment, operation, and maintenance of AI systems. This lack of transparency impedes inter-institutional learning and the development of reliable financial benchmarks ([Ibero-American Judicial Summit 2023](#)).

Activities

Once the problem has been identified, objectives defined, and available resources considered, it is necessary to detail a comprehensive portfolio of activities that will facilitate progress toward the desired outcomes. These activities must be articulated in a manner that is coherent with the specific institutional and regulatory context. While no single formula exists, the guidelines provided here are inspired by a series of established best practices and experiences available to date.

Conducting a Diagnostic Assessment

The conceptualization of the problems to be addressed through AI integration, combined with the institutional reflection processes described previously, will enable the formulation of a formal diagnostic document. Specific activities that will provide greater robustness to this diagnosis include, for example:

- Identification of technical and operational requirements.
- Assessment of data status and quality (data readiness).
- Investigation of implementation costs, for instance, through benchmarking with other judiciaries or by requesting budgetary quotes from external providers.
- Assessment of organizational change readiness and digital competencies.
- Evaluation of the available technological infrastructure for operation.
- Analysis of the applicable regulatory framework (data protection, transparency, fundamental guarantees in the judicial process, etc.).
- Digital security and cybersecurity analysis.
- Workflow analysis to identify bottlenecks and major opportunities for improvement.
- Mapping of institutional capabilities.
- Stakeholder mapping.
- Identification of specific risks and challenges by jurisdiction or operational area.

Establishment of Principles for the Implementation and Use of AI in the Judiciary

Several documents, at both the national and international levels, have formulated guiding principles for the regulation of AI in the justice sector or have addressed them when discussing its applications. These principles offer frameworks for institutional policy decisions and provide guidance regarding the primary dilemmas and decisions that must be adopted. Table 1 presents a brief systematization of the principles identified¹⁰, ordered by their frequency of occurrence. As can be observed, there is a high degree of convergence and coherence among the distinct sources.

Various national and international instruments have established guiding principles for the regulation of artificial intelligence (AI) in the justice sector or have articulated such principles when addressing its use. These principles provide a normative framework to inform institutional policymaking and serve as a reference for addressing the key dilemmas and decisions arising from the adoption of AI technologies. Table 1 presents a concise systematization of the principles identified across the reviewed sources, organized according to their frequency of occurrence. As shown, there is a high degree of convergence and consistency among the different frameworks.

Table 1. Principles for the Integration of Artificial Intelligence.

Principle	Countries or Documents
Human control and non-substitution of human rationality	Argentina (Chubut); Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; UNESCO; UN Special Rapporteur; Ibero-American Code of Ethics
Primacy of fundamental rights	Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; UNESCO; UN Special Rapporteur; Ibero-American Code of Ethics

¹⁰The documents consulted are identical to those detailed in the "Definition of Objectives" section.

Non-discrimination and prevention of bias	Argentina (Chubut); Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; New Zealand; UNESCO; UN Special Rapporteur; Ibero-American Code of Ethics
Transparency and explainability	Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; UNESCO; Ibero-American Code of Ethics
Accountability and responsibility	Argentina (Chubut); Colombia; Brazil; Argentina (CABA); New Zealand; UNESCO; Ibero-American Code of Ethics
Auditability and traceability	Brazil; Argentina (CABA); Spain; CEPEJ; UNESCO; Ibero-American Code of Ethics
Training, education, and informed use	Argentina (Chubut); Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; Canada; UNESCO
Impartiality, equality, and procedural fairness	Argentina (Chubut); Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; UNESCO; UN Special Rapporteur; Ibero-American Code of Ethics
Personal data protection and privacy	Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; New Zealand; UNESCO; Ibero-American Code of Ethics
Information security	Colombia; Brazil; Argentina (CABA); Spain; CEPEJ; UNESCO; Ibero-American Code of Ethics
Collaborative governance	Colombia; Brazil; Spain; CEPEJ; UNESCO; UN Special Rapporteur
Proportionality	Argentina (CABA); UNESCO; Ibero-American Code of Ethics
Due process and procedural guarantees	Brazil; Spain; CEPEJ; UN Special Rapporteur
Protection of Judicial Independence	Canada; UN Special Rapporteur

Source: author based on the guidelines and documents cited.

Formalization of Governance

The formalization of governance mandates that the rules regulating AI be enshrined within an official instrument, placing central emphasis on the allocation of competencies. This process organizes institutional management under a rigorous framework: identifying the locus of authority, establishing control standards, and defining accountability for outcomes. Regarding authority, it is

essential to formalize specific roles—such as Governance Committees or Chief AI Officers—to ensure that every project receives appropriate validation.

Formalizing governance also facilitates the establishment of mandatory operating protocols, standardizing the project life cycle and the risk categorization process. Finally, the governance of accountability must define the requisite levels of human oversight and process traceability, thereby clearly identifying liability in the event of system failures.

The OECD, in its report "[Governing with Artificial Intelligence](#)," specifically in the section regarding applications in the administration of justice¹¹, proposes adopting a governance approach that spans the entire life cycle of AI systems. This includes promoting continuous evaluation processes, technical audits, clear documentation, and sustained control mechanisms. Similarly, UNESCO's guidelines ([UNESCO 2025](#)) for the use of AI systems in courts and tribunals emphasize the necessity of establishing internal governance frameworks with clearly defined responsibilities and interdisciplinary oversight committees.

In alignment with these standards, the [National Institute of Standards and Technology \(NIST\) framework on AI risk management](#) proposes a structure based on four key functions for the responsible and secure management of automated systems:

- **Govern:** Cultivate a culture of risk management through clear institutional roles, responsibilities, and policies.
- **Map:** Contextualize the system to understand its operations, application environment, and potential impacts.
- **Measure:** Assess risks, biases, and performance using quantitative and verifiable metrics.
- **Manage:** Implement risk mitigation strategies, continuous monitoring, and adaptive capacity.

The development of this framework represents an opportunity to foster broad, participatory, and representative institutional reflection that incorporates the diverse voices, functions, and expertise of coexisting within the judiciary.

¹¹ Page 265.

Risk Stratification: Defining Use Categories

A strategic and essential step is establishing a framework that enables the classification of risks posed by implementing AI tools to fundamental rights. This requires risk stratification—creating a scale ranging from prohibited uses to those posing minimal risk—and providing clear definitions and examples for each category. Consequently, new tools will be categorized at each level, entailing specific operational implications in accordance with established guidelines, user manuals, and protocols.

The established principles serve as a guiding framework for this stratification process. This consistency between strategic objectives and risk levels enables the application of proportionality criteria in management: the higher the identified risk level, the more stringent the requirements for validation, auditing, and human oversight. The implementation of pilot plans and the evaluation of algorithmic impact play a pivotal role in this process, as detailed below.

This risk stratification approach is inspired by the European Union Artificial Intelligence Act ([EU AI Act](#)) of 2024, the first comprehensive global regulatory framework aimed at governing the development, marketing, and use of AI systems. Its provisions serve as a mandatory reference for drafting internal guidelines or protocols for AI governance in the justice sector. Indeed, this model has been adopted by the [Agreement of the Superior Council of the Judicature of Colombia](#), the regulations of the [National Council of Justice \(CNJ\) of Brazil](#), and the proposal for an [Ibero-American Code of Ethics for the Use of Artificial Intelligence](#). It establishes absolute prohibitions and sets strict requirements for "high-risk" systems, including conformity assessments, meaningful human oversight, transparency, traceability, and risk management mechanisms.

Development of Pilot Implementation Plans

Another recommended activity is the design and execution of pilot programs prior to full-scale deployment. These initiatives allow for the testing of tools in real-world applications, facilitating the

identification and management of operational risks and generating valuable lessons for the go/no-go decision¹².

Each pilot must be defined with:

- Clear operational objectives.
- Limited scope: Restricted functionality (single application) and institutional (single court or unit).
- Detailed deployment planning, concluding with an Algorithmic Impact Assessment (AIA).

This phase allows for the anticipation of risks and impacts generated by AI usage. The pilot stage serves as a decision gate. Based on the evidence gathered, a determination will be made regarding whether the tool can be implemented, requires redesign or adjustments, or if its deployment is inadvisable. This mechanism is particularly useful in cases where ambiguity exists regarding the risk level assigned to a specific application^{13 14}.

Algorithmic Impact Assessment and Ex-Post Evaluation¹⁵

Algorithmic Impact Assessment and Algorithmic Audits and Impact Evaluations are complementary tools that form part of a comprehensive approach to managing the risks associated with the use of AI in the administration of justice. Both processes focus on identifying, analyzing, and mitigating the adverse effects that an AI system may have on fundamental rights, institutional principles, specific populations, and the overall functioning of an organization

¹² The European Union AI Act promotes the establishment of regulatory sandboxes, defined as secure environments that permit developers to test AI systems under real-world conditions subject to strict regulatory oversight. These mechanisms facilitate the detection of failures, biases, or adverse impacts prior to deployment, particularly regarding high-risk applications.

¹³ The "[Pilot Program for the Strategic and Responsible Use of Generative AI in the Argentine Justice System](#)" conducted trials utilizing commercial generative AI tools (ChatGPT, Gemini, and Copilot) to support the drafting and formulation of adjudicative decisions.

¹⁴ The [Mikuna AI](#) pilot tests an AI-based platform designed to expedite the processing of child support claims. This system automates the analysis of judicial case files, recognizes and extracts key information from documents, structures data regarding the parties, and streamlines document management.

¹⁵ This distinction between analysis and evaluation is adopted from the draft UNESCO Guidelines for the Use of AI Systems in Courts in Tribunals ([UNESCO 2025](#)).

([UNESCO 2025](#)). However, they differ in terms of when they are applied and their specific purpose within the system's life cycle.

The Algorithmic Impact Assessment is carried out during the pilot or preliminary design phase, before the system goes into operation. Its objective is to anticipate possible impacts, such as likely positive and negative effects, or mapping potential risks like system failures. It allows parameters to be adjusted, early risks to be detected, and the tool's performance to be evaluated under controlled real-world conditions, in order to then decide on its possible final implementation. This analysis is essential to assess whether a given tool is technically feasible, socially acceptable, and legally compatible with the institutional context in which it is intended to be applied, specifically identifying potential impacts on human rights such as fair trial, due process, and access to justice ([Reisman et al. 2018](#)).

Once AI systems or tools are in operation, Algorithmic Audits and Impact Evaluations allow for the monitoring of the system's performance, checking whether biases not contemplated in the design emerge, or whether the tool continues to comply with regulations regarding fairness, transparency, and security. This process allows for the correction of deviations and ensures that the tool continues to comply with the required ethical and legal standards, evaluating the effects on users and affected populations. The judiciary should explicitly establish processes for third parties or workers to report potential vulnerabilities and risks that arise during this stage.

In both the assessment and the evaluation, the process usually begins with a structured methodology that inquires about the design of the system and the data used. This model is reflected in tools such as the Canadian government's Algorithmic Impact Assessment Tool, which classifies the risk level based on the system's characteristics. This classification guides the requirements, mitigation measures, or oversight mechanisms needed, ensuring that systems with greater potential for impact are subject to stricter controls.

In an assessment, it is key to define the functional purpose of the system and whether the process to be implemented is compatible with the principles that inform the justice system and procedural guarantees. This stage should also include a detailed analysis of the data, ensuring that the system allows for human intervention and that the judiciary maintains control over the data. For Generative

AI models, specific risks must be analyzed, such as the potential for generating incorrect, imprecise, or fictitious information (often referred to as hallucinations). It is critical to verify outputs against reliable sources and ensure that the system does not replace expert witness testimonies or independent legal analysis.

The ex-post evaluation, for its part, allows us to observe how the risks initially identified evolve and to detect new threats that arise with sustained use of the system. At this point, it is recommended to assess the need for technical adjustments or updates to the system's risk management framework.

Both processes should conclude with the definition of mitigation measures proportional to the risk. In the most extreme cases, if an assessment or evaluation finds that specific AI systems violate human rights or produce disproportionate negative impacts, the judiciary must refrain from adopting them or cease their use immediately and decommission the tool.

Finally, it is recommended that the results of both the assessments and the evaluations be published to reinforce transparency, accountability, and institutional legitimacy. This provides a key opportunity to encourage the involvement of civil society and independent auditors, who can contribute to accumulating knowledge and improving the administration of justice through external oversight.

Development of Guidelines, User Manuals, and Protocols

Based on the evidence reviewed, it is recommended to develop ethical frameworks, user manuals, and internal protocols for the use of AI in the judiciary. These instruments will explicitly codify the practices outlined above, ensuring their clear application in the design, supervision, and responsible use of these technologies.

The formulation of these documents must ensure the visibility and active participation of all levels of the judiciary, including those performing adjudicative functions, legal advisory, administrative management, user support, and technical support. These instruments will provide practical criteria indicating the valid contexts for use, the minimum safeguards required, and the necessary controls—effectively establishing the risk stratification process and its operational consequences.

Furthermore, they will provide clear operational instructions for the use of tools, result of validation, error detection, and the management of unreliable outputs. It is also recommended that these instruments establish periodic review processes, oversight mechanisms, feedback channels, and specific consequences associated with negligent use or misuse.

Implementation of Capacity Building and AI Literacy Programs

Another critically recommended activity is the systematic development of capacity building, awareness, and training activities tailored to different user profiles and levels of responsibility. Training needs are heterogeneous: while some personnel require basic technical skills to operate specific tools, others must develop AI literacy—specifically, a critical understanding of the limitations, biases, and risks associated with AI systems.

In this process, it is advisable to involve human resource managers to review job profiles, training plans, and existing institutional competencies. This will help identify gaps between current capabilities and the skills demanded by AI in daily operations. This capacity diagnosis will enable the planning of effective training interventions, the definition of technical support needs, and, if necessary, the redesign or creation of institutional roles to support digital transformation from a holistic perspective.

Simultaneously, training must raise awareness regarding the impacts of AI on adjudicative work, the fundamental rights of users, and the institutional culture. Training is therefore key to promoting responsible appropriation, preventing misuse, and fostering a critical attitude that enables risk identification and continuous improvement of these tools.

Rigorous Regulation of Training Data and Bias Mitigation

The absence of bias in training data is one of the most complex challenges in developing prediction and decision-support tools for the judiciary. The administration of justice relies on the principles of legality, impartiality, and equality before the law, requiring rigorous attention to the characteristics, origin, and validity of the data used.

While sample representativeness is frequently addressed, judiciaries must scrutinize additional factors: the age of the data, its normative validity (legal relevance), the institutional context of its generation, and the framework of guarantees in force at the time of production. Using large volumes of data does not guarantee bias elimination. Biases may stem from decisions based on repealed regulations, discriminatory social practices, or jurisprudential criteria abandoned by current legal practices. Furthermore, data generated in historical contexts lacking judicial independence or under authoritarian regimes pose significant risks. Training models with such data without critical review or corrective measures can reproduce historical biases, consolidate inequalities, and revalidate decisions lacking democratic legitimacy. Consequently, constructing predictive systems requires technical and methodological definitions made with the highest degree of reflection, transparency, and institutional deliberation.

Development of Performance and Impact Indicators

It is suggested to verify compliance with objectives through operational efficiency indicators, such as effective time reduction. Similarly, it is recommended to quantify actions performed with new tools (e.g., number of cases resolved via automated channels) and to develop evaluation instruments, such as user satisfaction surveys regarding these services.

Additionally, it is recommended to develop a Public Institutional Repository (or Register of Algorithms) to consolidate activities and centralize the catalog of deployed tools. This promotes transparency, replicability, and effective control over implemented developments. Finally, it is proposed to generate indicators for training and competency development, evaluating the actual coverage of digital and AI literacy among staff, as well as the rate of effective adoption of new tools in daily workflows.

Expected outcomes

The medium- and long-term benefits and the magnitude of the transformations implemented depend directly on the soundness of the implementation strategy adopted. The accuracy in identifying problems, the adequacy of the resources mobilized, and the relevance of the activities carried out will determine the real possibility of achieving the desired outcomes. These factors will inevitably condition the choice of certain uses of AI and the scope of the improvements they can bring to the justice service.

Examples of outcomes include:

- Increased public confidence and strengthened institutional legitimacy achieved through faster, more transparent, and more accountable case management.
- Enhanced strategic decision-making processes driven by the availability of more and better operational data.
- Improved access to justice facilitated by new tools that bring the justice system closer to the people.
- Consolidation of a modern and secure management model in which personnel are oriented toward the production of substantive value.

Conclusions

Artificial intelligence is profoundly transforming various sectors of modern society, and the administration of justice is no exception. In a landscape defined by rapid change and uncertainty regarding long-term effects, this report advocates for the adoption of these tools in a strategic, planned, and responsible manner, aiming to maximize benefits while minimizing risks to the fundamental rights of those accessing the justice system.

To achieve this, it is proposed that judiciaries utilize a Theory of Change framework. This approach begins with a clear identification of the public problems to be resolved, followed by an accurate diagnosis to guide resource allocation and the definition of activities and products that will effectively achieve the desired outcomes and impacts.

In this process, the tension between efficiency gains and the protection of rights must remain central to the debate. To this end, the distinction between administrative management and adjudicative functions is fundamental. While the automation of administrative tasks offers a secure, low-risk avenue to address many structural deficiencies in judicial services, the incorporation of these tools into the adjudicative function requires extreme caution. In this context, distinguishing between support tools and autonomous or substitutive systems is critical, as the latter threaten the very essence of the judicial function.

This technological inflection point offers the region's judiciaries a renewed opportunity—perhaps the last—to strengthen the administration of justice and restore citizens' eroding confidence in the judicial system. A responsible and planned implementation can yield significant gains in legitimacy, access, and efficiency, paving the way for a new stage of institutional consolidation. Conversely, unregulated and ad hoc deployment may lead to undesirable, permanent consequences: raising unprecedented questions about the necessity of human justice, fueling narratives advocating for the total displacement of judges, and ultimately causing the judiciary to lose its relevance as an essential counterweight to political power in a democratic state governed by the rule of law.

References

Alarie, B. (2016, 27 de mayo). *The path of the law: Toward legal singularity*. SSRN. <https://ssrn.com/abstract=2767835>

Alarie, B., Niblett, A., & Yoon, A. H. (2018). How artificial intelligence will affect the practice of law. *University of Toronto Law Journal*, 68(Suppl. 1), 106–124. <https://www.jstor.org/stable/90019655>

Buchholtz, G. (2020). Artificial intelligence and legal tech: Challenges to the rule of law. In T. Wischmeyer & T. Rademacher (Eds.), *Regulating artificial intelligence*. Springer. https://doi.org/10.1007/978-3-030-32361-5_8

Castets-Renard, C. (2021). Human rights and algorithmic impact assessment for predictive policing. In H.-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, & G. De Gregorio (Eds.), *Constitutional challenges in the algorithmic society* (pp. 93–110). Cambridge University Press.

Chiao, V. (2019). Fairness, accountability and transparency: Notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, 15(2), 126–139. <https://doi.org/10.1017/S1744552319000077>

Cofone, I., & Khern-am-nuai, W. (2025). The overstated cost of AI fairness in criminal justice. *Indiana Law Journal*, 100(4), Article 4. <https://www.repository.law.indiana.edu/ilj/vol100/iss4/4>

Cohen, M. C., Dahan, S., & Khern-am-nuai, W. (2023). The use of AI in legal systems: Determining independent contractor vs. employee status. *Artificial Intelligence and Law*. <https://doi.org/10.1007/s10506-023-09353-y>

Coglianesi, C., & Lehr, D. (2019). Transparency and algorithmic governance. *Administrative Law Review*, 71, 1–?. <https://ssrn.com/abstract=3293008>

Dias, S. A. D. J., & Sátiro, R. M. (2024). Artificial intelligence in the judiciary: A critical view. *Futures*, 164, Article 103493. <https://doi.org/10.1016/j.futures.2024.103493>

El Mestari, S. Z., Lenzini, G., & Demirci, H. (2024). Preserving data privacy in machine learning systems. *Computers & Security*, 137, Article 103605. <https://doi.org/10.1016/j.cose.2023.103605>

Ferrara, E. (2023, 27 de octubre). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. SSRN. <https://ssrn.com/abstract=4615421>

Gutiérrez, J. D. (2024). Critical appraisal of large language models in judicial decision-making. In R. Paul, E. Carmel, & J. Cobbe (Eds.), *Handbook on public policy and artificial intelligence* (pp. 323–338). Edward Elgar Publishing. <https://doi.org/10.4337/9781803922171.00033>

Hooshidary, S., Canada, C., & Clark, W. (2025). *Artificial intelligence in government: The federal and state landscape*. National Conference of State Legislatures.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, 133(1), 237–293. <https://doi.org/10.1093/qje/qjx032>

Kolkman, D., Bex, F., Narayan, N., & van der Put, M. (2024). Justitia ex machina: The impact of an AI system on legal decision-making and discretionary authority. *Big Data & Society*, 11(2). <https://doi.org/10.1177/20539517241255101>

Malek, M. A. (2022). Criminal courts' artificial intelligence: The way it reinforces bias and discrimination. *AI Ethics*, 2, 233–245. <https://doi.org/10.1007/s43681-022-00137-9>

McGregor, L., Murray, D., & Ng, V. (2019). International human rights law as a framework for algorithmic accountability. *International and Comparative Law Quarterly*, 68(2), 309–343. <https://doi.org/10.1017/S0020589319000046>

Pérez-Pacheco, Y., & Simón Castellano, P. (2025). La inteligencia artificial en los sistemas judiciales de América Latina: Una revisión sobre desafíos y oportunidades. *A&C – Revista de Direito Administrativo & Constitucional*, 25(100), 95–129. <https://doi.org/10.21056/aec.v25i100.2027>

Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). *Algorithmic impact assessments: A practical framework for public agency accountability*. AI Now Institute.

Ribeiro, M. R. B., & Carneiro, G. de F. (2025). An exploratory study of the impact of artificial intelligence on the judiciary productivity and efficiency. *ARACÊ*, 7(4), 18104–18119. <https://doi.org/10.56238/arev7n4-144>

UNESCO. (2023). *Kit de herramientas sobre la IA y el Estado de derecho*. UNESCO. <http://www.unesco.org/open-access/terms-use-ccbysa-en>

Vargas-Murillo, A. R., de la Asunción Pari-Bedoya, I. N. M., Turriate-Guzmán, A. M., Delgado-Chávez, C. A., & Sánchez-Paucar, F. (2024). Transforming justice: Implications of artificial intelligence in legal systems. *Academic Journal of Interdisciplinary Studies*, 13(2), 433. <https://doi.org/10.36941/ajis-2024-0059>

West, S. M. (2020). Redistribution and Rekognition: A feminist critique of algorithmic fairness. *Catalyst: Feminism, Theory, Technoscience*, 6(2), 1–24. <https://doi.org/10.28968/cftt.v6i2.33613>

Yang, X., Wang, Z., Wang, Q., Wei, K., Zhang, K., & Shi, J. (2024). Large language models for automated Q&A involving legal documents: A survey on algorithms, frameworks and applications. *International Journal of Web Information Systems*, 20(4), 413–435. <https://doi.org/10.1108/IJWIS-12-2023-0256>



Centro de Estudios de Justicia de las Américas
Justice Studies Center of the Americas

Artificial Intelligence in the Judiciary: Reflections and Guidelines for the Americas



Centro de Estudios de Justicia de las Américas (CEJA)

www.cejamericas.org

