

La actividad Volitiva del Juez Penal en la era de la Inteligencia Artificial: Reflexiones sobre algoritmos y Justicia Predictiva

The Volitional activity of the Criminal Judge in the age of Artificial Intelligence: Reflections on algorithms and Predictive Justice

José María Caballero Galeano¹

RESUMEN

La irrupción de la inteligencia artificial (IA) en el ámbito del derecho penal plantea interrogantes fundamentales sobre la autonomía y la actividad volitiva del juez. Sistemas como COMPAS en Estados Unidos y HART en el Reino Unido utilizan algoritmos predictivos para evaluar riesgos y apoyar decisiones judiciales. Sin embargo, estas herramientas, al igual que la tecnología ficticia de la película *Minority Report*, anticipan escenarios futuros basados en patrones, cuestionando el principio de presunción de inocencia y las garantías procesales. Este artículo explora la analogía entre los algoritmos y la película, analizando los sesgos inherentes a estas tecnologías y los posibles perjuicios del uso de herramientas como el reconocimiento facial. Finalmente, se reflexiona sobre la responsabilidad ética y jurídica: ¿De quién es la culpa de los errores? ¿Del juez, de la IA o de los programadores?

Palabras clave: Inteligencia Artificial, Derecho Penal, Justicia Predictiva, Sesgos Algoritmos, Reconocimiento Facial.

ABSTRACT

The emergence of artificial intelligence (AI) in the field of criminal law raises fundamental questions about the autonomy and volitional activity of the judge. Systems such as COMPAS in the United States and HART in the United Kingdom use predictive algorithms to assess risks and support judicial

¹ CABALLERO GALEANO, José María. Doctor en Derecho Público y Doctor en Ciencias Jurídicas por la Universidad Columbia del Paraguay, con estudios posdoctorales en Derecho en la Universidad de Bolonia, Italia. Es Magíster en Derecho Penal y Procesal Penal, Ciencia Política y Ciencias de la Educación. Abogado egresado de la UNA, con múltiples especializaciones en investigación, gobernabilidad, justicia constitucional y docencia universitaria. Actualmente se desempeña como Decano de la Facultad de Ciencias Jurídicas, Políticas y Sociales de la Universidad Autónoma de Asunción, y es docente de posgrado, tutor y miembro de mesa examinadora en diversas maestrías y doctorados. Egresado de la Escuela Judicial y del Centro de Estudios Hemisféricos de Defensa William J. Perry, es un referente en temas de defensa, gobernanza y educación superior.

decisions. However, these tools, like the fictional technology in the movie *Minority Report*, anticipate future scenarios based on patterns, calling into question the principle of presumption of innocence and due process. This article explores the analogy between algorithms and the movie, analyzing the biases inherent in these technologies and the potential harms of using tools such as facial recognition. Finally, it reflects on the ethical and legal responsibility: whose fault are the errors: the judge's, the AI's or the programmers'?

Keywords: Artificial Intelligence, Criminal Law, Predictive Justice, Bias Algorithms, Facial Recognition.

Introducción

La inteligencia artificial (IA) ha emergido como una de las fuerzas tecnológicas más transformadoras de nuestro tiempo, redefiniendo la manera en que las sociedades abordan problemas complejos, toman decisiones y gestionan información. En el ámbito jurídico, y particularmente en el derecho penal, la IA ha introducido herramientas que prometen revolucionar la administración de justicia al ofrecer soluciones basadas en el análisis masivo de datos, la identificación de patrones ocultos y la predicción de comportamientos futuros. Estas capacidades, inéditas en la historia del derecho, representan un cambio de paradigma en cómo se entienden los procesos judiciales, la toma de decisiones y el rol de los actores que participan en ellos, especialmente el juez penal.

Dentro de este contexto, destacan herramientas como COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) y HART (Harm Assessment Risk Tool), diseñadas para asistir a los jueces en la evaluación de riesgos de reincidencia y peligrosidad. Estas tecnologías, que combinan estadística avanzada y aprendizaje automático, prometen mejorar la eficiencia y precisión en decisiones judiciales críticas, como la imposición de medidas cautelares o la concesión de libertad condicional. Sin embargo, su incorporación plantea cuestiones fundamentales que trascienden lo técnico y abordan los valores centrales de la justicia penal: ¿cómo garantizar que estas herramientas no perpetúen desigualdades sociales o discriminen a ciertos grupos? ¿Hasta qué punto es prudente delegar en algoritmos decisiones que afectan derechos y libertades fundamentales?

Además, el impacto cultural y filosófico de estas tecnologías no puede ser ignorado. La película *Minority Report* ofrece una narrativa distópica que, aunque ficticia, resuena en el contexto actual. En ella, un sistema de justicia preventiva se basa en predicciones realizadas por "precogs", capaces de prever crímenes antes de que ocurran. Aunque los algoritmos modernos no predicen crímenes con la certeza casi mágica de esta ficción, sí intentan proyectar probabilidades basadas en patrones del pasado, lo que introduce un debate ético crucial: ¿es justo anticipar responsabilidades penales en función de predicciones estadísticas? ¿Cómo afecta esto al principio de presunción de inocencia, piedra angular del derecho penal?

En este marco, el rol del juez penal adquiere una relevancia singular. Históricamente, el juez ha sido el intérprete de la ley, el evaluador de las pruebas y el garante de los derechos procesales de las partes. Su actividad volitiva, entendida como la capacidad de tomar decisiones fundamentadas en un juicio ético y humano, está siendo reconfigurada por la presencia de estas tecnologías. El riesgo no es menor: la dependencia excesiva en herramientas algorítmicas podría convertir al juez en un mero validador de decisiones tecnológicas, debilitando su autonomía y de humanizando el proceso penal.

Por otro lado, los algoritmos no son entes neutrales. Están diseñados y entrenados con datos que reflejan las desigualdades históricas y estructurales de la sociedad. Por ejemplo, estudios realizados sobre COMPAS han demostrado que este sistema tiende a sobreestimar el riesgo de reincidencia en personas afroamericanas y a subestimarlo en personas blancas. Este fenómeno, conocido como sesgo algorítmico, no es accidental, sino el resultado de un diseño que replica las dinámicas de exclusión y discriminación presentes en los datos de entrenamiento. Así, en lugar de corregir desigualdades, estas tecnologías pueden reforzarlas, socavando la confianza pública en el sistema de justicia.

Finalmente, el contexto tecnológico del derecho penal no puede entenderse sin reconocer su impacto en los principios fundamentales del sistema jurídico. La incorporación de la IA no solo plantea retos técnicos, como la interpretación de resultados algorítmicos o la transparencia en su funcionamiento, sino también dilemas éticos y jurídicos de gran calado. ¿Debe la tecnología ser un complemento o un sustituto del juicio humano? ¿Cómo garantizar que la eficiencia no sacrifique los valores de equidad y justicia? Estas preguntas configuran el trasfondo del debate contemporáneo sobre el uso de la IA en el derecho penal, un debate

que requiere un análisis profundo y multidisciplinario para preservar el equilibrio entre innovación tecnológica y respeto por los derechos humanos.

La introducción de la inteligencia artificial (IA) en el ámbito del derecho penal representa mucho más que un avance técnico; es un desafío profundo que redefine conceptos fundamentales de justicia, equidad y autonomía judicial. En el corazón de este cambio se encuentra una pregunta esencial: ¿cómo puede la tecnología, con su capacidad de procesar datos masivos y ofrecer predicciones estadísticas, integrarse en un sistema jurídico diseñado para valorar la individualidad, la responsabilidad y los derechos humanos? La incorporación de herramientas como COMPAS y HART, diseñadas para evaluar riesgos de reincidencia y peligrosidad, pone de manifiesto tanto las promesas como las amenazas de la IA en el contexto penal.

Estas tecnologías, aunque sofisticadas, no están exentas de limitaciones. Los algoritmos operan sobre la base de patrones extraídos de datos históricos, que a menudo reflejan desigualdades estructurales y prejuicios sociales. Por ejemplo, diversos estudios han documentado cómo COMPAS tiende a producir evaluaciones sesgadas que sobreestiman los riesgos asociados a individuos afroamericanos y subestiman los de personas blancas. Esto plantea un dilema ético fundamental: ¿puede un sistema de justicia considerado imparcial depender de herramientas que, aunque eficaces, reproducen sesgos históricos? ¿Cómo garantizar que la tecnología no se convierta en una herramienta para perpetuar desigualdades, sino en un medio para superarlas?

Otro aspecto crítico es el impacto de la IA en la actividad volitiva del juez penal. Tradicionalmente, el juez ha sido la figura central en el proceso penal, ejerciendo un juicio ético y humano que combina conocimiento jurídico, experiencia profesional y sensibilidad contextual. La llegada de la IA introduce el riesgo de una excesiva delegación de responsabilidades en sistemas automatizados, lo que podría reducir al juez a un mero validador de resultados algorítmicos. Este desplazamiento no solo compromete la autonomía del juez, sino que también transforma la naturaleza misma del proceso penal, llevándolo hacia una lógica tecnocrática que podría deshumanizar la justicia.

El principio de presunción de inocencia, pilar del derecho penal, también enfrenta nuevos desafíos en la era de la IA. Herramientas como HART y COMPAS, que proyectan

probabilidades de reincidencia basadas en datos pasados, transforman la valoración de la peligrosidad en un ejercicio de predicción estadística. Esta lógica, aunque útil para ciertos fines administrativos, entra en conflicto con la idea de que las personas solo pueden ser responsabilizadas penalmente por actos efectivamente cometidos y no por riesgos o probabilidades futuras. Aquí surge una pregunta clave: ¿hasta qué punto es legítimo que el sistema penal actúe sobre la base de predicciones, sacrificando principios tradicionales en nombre de la eficiencia?

En este contexto, las narrativas culturales y filosóficas ayudan a iluminar los dilemas que plantea la IA. La película *Minority Report*, por ejemplo, presenta un sistema de justicia basado en la previsión de crímenes antes de que ocurran, administrado por "precogs" que determinan el futuro basándose en datos y percepciones. Aunque ficticio, este escenario ofrece un paralelismo inquietante con el uso actual de algoritmos en el proceso penal. En ambos casos, el riesgo principal radica en cómo las predicciones pueden desvirtuar la concepción de justicia y llevar a decisiones profundamente injustas.

Este artículo aborda las implicaciones de la IA en el derecho penal a través de tres ejes principales. En primer lugar, analiza cómo las herramientas algorítmicas afectan la autonomía judicial y la actividad volitiva del juez, cuestionando su rol tradicional como garante de la justicia y los derechos procesales. En segundo lugar, examina los riesgos asociados a la justicia predictiva, incluyendo los sesgos algorítmicos y las limitaciones inherentes al aprendizaje automático. Finalmente, explora cómo la IA redefine los principios básicos del derecho penal, como la presunción de inocencia y la responsabilidad individual, abriendo el debate sobre los límites éticos y legales de estas tecnologías.

La integración de la IA en el sistema penal no es solo una cuestión de innovación tecnológica, sino un reto multidimensional que exige reflexiones profundas y soluciones equilibradas. Este análisis busca ofrecer una visión crítica sobre cómo preservar los valores fundamentales del derecho penal en una era marcada por cambios tecnológicos sin precedentes, destacando la necesidad de equilibrar progreso y principios en la administración de justicia.

Desarrollo

La inteligencia artificial (IA) está reconfigurando de manera significativa el rol del juez penal y el desarrollo del proceso judicial, particularmente a través de la implementación de

herramientas algorítmicas como COMPAS y HART. Tradicionalmente, el juez penal ha desempeñado una función central como garante de los derechos fundamentales y árbitro imparcial, ejerciendo su actividad volitiva para interpretar la ley y valorar las pruebas en contextos éticos y humanos específicos. Sin embargo, la introducción de estas tecnologías plantea el riesgo de que esta actividad volitiva sea reemplazada o condicionada por decisiones automatizadas, lo que podría deshumanizar el sistema de justicia penal.

COMPAS y HART, diseñados para evaluar riesgos de reincidencia y peligrosidad, son herramientas que prometen ofrecer predicciones rápidas y precisas basadas en el análisis de grandes volúmenes de datos. Sin embargo, estas promesas se ven empañadas por evidentes limitaciones, particularmente en lo que respecta a la transparencia y la imparcialidad. Un estudio de ProPublica sobre COMPAS reveló que este sistema sobreestima sistemáticamente el riesgo de reincidencia en personas afroamericanas y lo subestima en personas blancas. Este sesgo algorítmico no es fortuito; es el resultado de entrenar algoritmos con datos históricos que reflejan desigualdades estructurales de larga data. En lugar de corregir estas desigualdades, las herramientas algorítmicas a menudo las perpetúan y amplifican, generando decisiones que parecen objetivas pero que están profundamente influenciadas por dinámicas discriminatorias.

Cuando los jueces dependen excesivamente de estas herramientas, se enfrentan al riesgo de caer en lo que se conoce como automatización del sesgo: una tendencia a aceptar sin cuestionar las recomendaciones algorítmicas debido a su aparente objetividad. Esto debilita la autonomía judicial y transforma al juez en un mero validador de decisiones tecnológicas. La actividad volitiva del juez, entendida como su capacidad para tomar decisiones fundamentadas en principios éticos, jurídicos y contextuales, queda relegada en favor de proyecciones estadísticas que no consideran las complejidades individuales de cada caso. Este desplazamiento representa una amenaza no solo para la autonomía del juez, sino también para la naturaleza misma del proceso penal.

Uno de los principios más afectados por la justicia predictiva basada en IA es la presunción de inocencia. Este principio, pilar del derecho penal, establece que una persona es inocente hasta que se demuestre lo contrario mediante pruebas claras y contundentes de un acto cometido. Sin embargo, las herramientas algorítmicas cambian el enfoque hacia la predicción de comportamientos futuros basados en datos del pasado. Por ejemplo, un acusado puede

enfrentar restricciones de libertad no por actos cometidos, sino por proyecciones de que podría reincidir en el futuro. Este enfoque, aunque útil para fines administrativos, entra en conflicto con el principio de responsabilidad penal basado en hechos concretos y plantea una pregunta crítica: ¿hasta qué punto es legítimo actuar preventivamente basándose en predicciones estadísticas? ¿No transforma esto al sistema penal en una maquinaria de control social más que en un mecanismo de justicia?

El reconocimiento facial es otro ejemplo de cómo la tecnología puede generar tanto oportunidades como riesgos en el ámbito penal. Si bien estas herramientas han demostrado ser útiles para identificar sospechosos y acelerar investigaciones, su aplicación ha generado numerosas controversias debido a sus altos índices de error, especialmente en mujeres y personas de piel oscura. Estos errores, que se originan en bases de datos insuficientemente diversas y en diseños algorítmicos deficientes, han llevado a detenciones injustas y violaciones de derechos fundamentales. Este tipo de fallos no solo afecta a los individuos directamente implicados, sino que también socava la confianza pública en el sistema judicial y plantea interrogantes sobre quién debe asumir la responsabilidad por los errores tecnológicos: ¿el juez que confió en la herramienta, el programador que la diseñó o la institución judicial que permitió su implementación sin controles adecuados?

Además de estos problemas específicos, la dependencia en herramientas algorítmicas plantea un riesgo más amplio de deshumanización del proceso penal. El sistema judicial, diseñado para valorar las circunstancias únicas de cada caso y garantizar que se respeten los derechos de todas las partes, podría convertirse en un mecanismo tecnocrático donde las decisiones se basan exclusivamente en modelos estadísticos. Esto erosiona la percepción de la justicia como un proceso ético y humano, reduciendo a los acusados a meras cifras en un sistema automatizado. La figura del juez, en lugar de ser un garante de los derechos individuales y la imparcialidad, corre el riesgo de convertirse en un validador pasivo de decisiones tecnológicas.

La justicia penal no solo se enfrenta al desafío de incorporar herramientas tecnológicas con eficacia, sino también al reto de hacerlo sin comprometer los valores fundamentales del sistema. La actividad volitiva del juez, como expresión de su juicio ético y humano, debe ser preservada como un elemento central del proceso penal. Aunque la IA puede ofrecer información valiosa y complementar el trabajo judicial, no debe sustituir el criterio humano

ni desplazar los principios que sustentan el derecho penal. La tecnología debe ser una herramienta al servicio de la justicia, no un sustituto de sus valores fundamentales.

Este contexto pone de manifiesto la necesidad de una regulación clara y estricta que limite el impacto de los sesgos algorítmicos, garantice la transparencia de los sistemas de IA y preserve la autonomía de los jueces. Asimismo, se requiere una supervisión constante y multidisciplinaria que combine la experiencia de juristas, tecnólogos y especialistas en derechos humanos para asegurar que el uso de la IA en el ámbito penal se alinee con los principios de justicia, equidad y dignidad humana.

Conclusiones

La incorporación de la inteligencia artificial (IA) en el derecho penal representa una transformación de proporciones históricas que redefine los fundamentos de la justicia, la autonomía judicial y los derechos fundamentales. Si bien las herramientas algorítmicas como COMPAS y HART han introducido avances significativos en términos de eficiencia y capacidad para procesar grandes volúmenes de datos, sus implicaciones van más allá de lo técnico. Estas tecnologías desafían principios jurídicos esenciales, como la presunción de inocencia, la igualdad ante la ley y la autonomía del juez, planteando dilemas éticos, legales y sociales de gran calado.

La actividad volitiva del juez, entendida como su capacidad de tomar decisiones fundamentadas en un juicio ético y humano, está en el centro de este debate. La IA, al proporcionar predicciones estadísticas basadas en patrones históricos, corre el riesgo de transformar al juez en un mero validador de resultados algorítmicos. Este fenómeno no solo amenaza la autonomía judicial, sino que también deshumaniza el proceso penal, reduciendo las decisiones judiciales a cálculos numéricos que ignoran la complejidad y singularidad de cada caso. La justicia no puede depender únicamente de probabilidades estadísticas; debe ser el reflejo de valores humanos, principios éticos y una comprensión contextual que solo un juez capacitado puede ofrecer.

Uno de los aspectos más preocupantes de la IA en el derecho penal es su capacidad para perpetuar y amplificar desigualdades estructurales. Los sesgos algorítmicos, como los documentados en herramientas como COMPAS, demuestran que estas tecnologías no son neutrales. Los datos con los que se entrenan los algoritmos suelen reflejar las desigualdades

históricas de las sociedades, reproduciendo patrones de discriminación racial, económica y de género. Cuando estos sesgos se institucionalizan bajo la apariencia de objetividad tecnológica, no solo se socava la confianza en el sistema de justicia, sino que también se vulneran los derechos de los grupos más marginados, contraviniendo el principio de igualdad ante la ley.

El impacto de estas tecnologías en la presunción de inocencia es igualmente preocupante. Herramientas como HART, que predicen comportamientos futuros basándose en patrones del pasado, introducen un modelo de justicia preventiva que contradice los fundamentos del derecho penal. En lugar de juzgar a las personas por actos efectivamente cometidos, estas tecnologías las clasifican y sancionan en función de riesgos potenciales. Este enfoque no solo erosiona la base misma del derecho penal, sino que también abre la puerta a un sistema de control social donde las decisiones judiciales se basan en probabilidades estadísticas en lugar de pruebas concretas.

El reconocimiento facial, otra aplicación relevante de la IA en el derecho penal, agrava estas preocupaciones. Aunque útil en investigaciones, esta tecnología ha demostrado ser particularmente propensa a errores, especialmente en el reconocimiento de mujeres y personas de piel oscura. Estas fallas han llevado a detenciones injustas y violaciones graves de derechos humanos, cuestionando la ética de depender de herramientas imperfectas en decisiones que afectan profundamente la vida de las personas. La falta de transparencia y supervisión en la implementación de estas tecnologías solo exacerba el problema, dejando a los afectados sin recursos claros para cuestionar su uso o apelar decisiones basadas en ellas.

Frente a estos desafíos, es imperativo establecer un marco regulatorio robusto y ético para la implementación de la IA en el derecho penal. Este marco debe abordar la transparencia de los algoritmos, exigir auditorías constantes para detectar y corregir sesgos, y garantizar que las decisiones judiciales sigan siendo humanas, éticas y fundamentadas en el contexto de cada caso. La tecnología debe ser un complemento al juicio humano, no un sustituto. Los jueces deben mantener su papel central como garantes de los derechos fundamentales, evaluando críticamente las recomendaciones algorítmicas y asegurándose de que estas no comprometan los principios esenciales de justicia y equidad.

Además, la responsabilidad por los errores tecnológicos debe ser claramente asignada. Los desarrolladores de herramientas de IA, las instituciones que las implementan y los jueces que las utilizan tienen roles interdependientes en este ecosistema. Es crucial que existan mecanismos claros para garantizar la rendición de cuentas cuando las decisiones basadas en IA resultan en errores o injusticias. Esto no solo reforzará la confianza pública en el sistema de justicia, sino que también incentivará el desarrollo de tecnologías más justas, transparentes y fiables.

El desafío que plantea la IA en el derecho penal no debe verse únicamente como una amenaza, sino también como una oportunidad para transformar el sistema de justicia en uno más eficiente, accesible y equitativo. Sin embargo, esta transformación no debe realizarse a costa de los principios fundamentales que han sustentado el derecho penal durante siglos. La verdadera justicia no puede ser reducida a un conjunto de cálculos algorítmicos; debe reflejar los valores éticos, los derechos fundamentales y el compromiso con la dignidad humana que definen a una sociedad verdaderamente democrática.

En última instancia, el futuro de la justicia penal en la era de la IA dependerá de nuestra capacidad para equilibrar innovación tecnológica con los principios éticos y jurídicos que garantizan una justicia genuina. Este equilibrio solo será posible mediante un enfoque colaborativo que integre a juristas, tecnólogos, legisladores y defensores de derechos humanos. Juntos, deben trabajar para garantizar que la IA sea una herramienta al servicio de la justicia y no una amenaza para ella. Preservar la autonomía judicial, proteger los derechos fundamentales y garantizar la equidad en el uso de la tecnología no son solo desafíos del presente, sino compromisos ineludibles para el futuro de la justicia penal.

Bibliografía

- Agüero, L. & Taramasco, C. (2022). *Inteligencia artificial y sesgos algorítmicos en el derecho penal: Un análisis crítico*. Tirant lo Blanch.
- Atienza, M. (2021). *Los desafíos éticos y jurídicos de la inteligencia artificial en el derecho*. Marcial Pons.
- Bueno, C. & Cuadrado, I. (2021). *Los algoritmos en el proceso penal: Implicaciones constitucionales y retos futuros*. Bosch.

- Díaz, M. (2021). *Derecho penal y nuevas tecnologías: Los límites de la inteligencia artificial en la decisión judicial*. Dykinson.
- Escobar, J. (2020). *Inteligencia artificial y responsabilidad penal: Análisis jurídico y filosófico*. Editorial Jurídica Continental.
- Fernández, R. (2020). *Justicia predictiva y automatización judicial: Riesgos y oportunidades*. Reus.
- López, C. (2019). *El juez y la inteligencia artificial: ¿Aliado o amenaza?* Editorial Aranzadi.
- Rouvier, A. (2019). *Inteligencia artificial y derecho: Consideraciones sobre su impacto en la administración de justicia*. Ediciones Jurídicas Cuyo.
- Sanz, M. (2021). *Algoritmos, big data y derecho penal: ¿Hacia una justicia deshumanizada?* Editorial Comares.
- Vázquez, P. (2022). *Inteligencia artificial y sesgos algorítmicos en el proceso penal*. Iustel.