



Revista Iberoamericana de Derecho, Cultura y Ambiente



Edición N° 9 – Julio de 2026

Capítulo de Derecho Penal y Criminología

www.aidca.org/revista

IA: OPACIDAD ALGORÍTMICA, LEGALIDAD Y DEBIDO PROCESO

Por José Ignacio Díaz Vélez¹

Introducción

Las garantías constitucionales en el proceso penal constituyen una serie de derechos orientados a salvaguardar al individuo frente al ejercicio del poder punitivo del Estado. Entre ellas, el principio de legalidad y el debido proceso ocupan un lugar central: el primero no solo opera como límite al legislador en la definición de los tipos penales, sino que también se proyecta sobre la actividad investigativa y sobre los fundamentos de la imputación concreta; el segundo garantiza que todo el andamiaje procesal —la acusación, la prueba, la defensa y la decisión judicial— respete las condiciones mínimas de racionalidad y contradicción que legitiman el ejercicio del poder punitivo. Es en estas dimensiones donde las herramientas de inteligencia artificial —en adelante IA—

¹Abogado, Universidad Nacional de Tucumán. Especialista en Justicia Constitucional y Derechos Humanos, Universidad de Bolonia. Especialista en Derecho Penal, Universidad Nacional de Rosario. Empleado Juzgado Federal N° 1 de Tucumán.



plantean desafíos que el ordenamiento jurídico argentino aún no ha abordado de manera sistemática.

En los últimos años, el uso de sistemas de IA en la investigación penal ha avanzado de manera significativa. El reconocimiento facial automatizado, los algoritmos de evaluación de riesgo de reincidencia, la predicción del delito basada en datos históricos y el procesamiento masivo de información digital son herramientas que hoy integran —o aspiran a integrar— la rutina de las fuerzas de seguridad y del sistema de justicia penal en numerosos países. En la República Argentina, la reciente creación de la Unidad de Inteligencia Artificial Aplicada a la Seguridad (UIAAS) mediante la Resolución 710/2024 del Ministerio de Seguridad de la Nación refleja con claridad esa tendencia y hace urgente la discusión que este trabajo propone.

Ahora bien, la pregunta que interesa no es si deben incorporarse herramientas de IA al sistema judicial —lo que sucederá indefectiblemente— sino cómo debe ejecutarse para que sean compatibles con las garantías constitucionales y convencionales que rigen la persecución penal en un Estado de Derecho. La IA, utilizada como herramienta de apoyo a la labor investigativa, ofrece beneficios concretos: mayor velocidad en el análisis de grandes volúmenes de información, reducción de errores humanos en tareas de identificación, y optimización de recursos en la persecución de delitos complejos. Sin embargo, la IA no puede ser concebida como un investigador autónomo ni, menos aún, como un juzgador. El ser humano —el juez, el fiscal, el defensor— debe conservar en todo momento el control racional y jurídico del proceso.

El problema que este trabajo examina es específico: la opacidad algorítmica, es decir, la imposibilidad de comprender el razonamiento interno mediante el cual un sistema de IA llega a un resultado, lo cual genera tensiones estructurales con el principio de legalidad y debido proceso en las dimensiones que se analizarán en los párrafos siguientes. Cuando una imputación se construye —en todo o en parte— sobre la base de un resultado algorítmico que el imputado no puede conocer ni refutar, que la defensa no puede controvertir y que el juez no



puede verificar racionalmente, el proceso penal deja de satisfacer las exigencias mínimas del Estado constitucional de Derecho.

El objetivo de este trabajo consiste en analizar esas tensiones, identificar los estándares constitucionales y convencionales aplicables al derecho argentino, y proponer criterios mínimos que deben cumplir los sistemas de IA utilizados en la investigación penal para ser compatibles constitucionalmente. Para ello, se examinarán casos paradigmáticos del derecho comparado, la experiencia argentina más relevante en la materia y el marco regulatorio europeo como referencia de derecho comparado.

I. La inteligencia artificial en la investigación penal: descripción funcional y beneficios

Antes de abordar los problemas jurídicos que plantea el uso de IA en la investigación penal, resulta necesario precisar a qué nos referimos cuando hablamos de estas herramientas y cuáles son sus potenciales beneficios, sin los cuales el análisis jurídico quedaría descontextualizado.

El Reglamento de Inteligencia Artificial de la Unión Europea (AI Act)² define a los sistemas de IA en su artículo 3 como *"un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales."* Esta definición es relevante no solo como referencia normativa, sino porque pone en evidencia lo que distingue a los sistemas de IA de otros instrumentos tecnológicos: su capacidad de generar resultados de manera semi-autónoma a partir de un proceso que no siempre es transparente.

²Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (AI Act). Publicado en el Diario Oficial de la Unión Europea el 12 de julio de 2024.



En este sentido, los sistemas de IA relevantes para la investigación penal pueden clasificarse, a grandes rasgos, en tres categorías. En primer lugar, los sistemas de *reconocimiento e identificación biométrica*, que procesan imágenes de cámaras de vigilancia para identificar personas buscadas por la justicia o detectar actividades sospechosas en tiempo real. En segundo lugar, los sistemas de *evaluación de riesgo*, que analizan datos históricos de una persona —historial delictivo, condiciones socioeconómicas, vínculos sociales, entre otros— para generar una puntuación que pretende predecir la probabilidad de que cometa un delito o reincida en uno. En tercer lugar, los sistemas de *análisis predictivo del delito* (*predictive policing*), que procesan datos georreferenciados, históricos y estadísticos para identificar zonas o grupos de riesgo donde concentrar la actividad preventiva.

Cada una de estas herramientas ofrece ventajas concretas como instrumento de apoyo al trabajo humano. Un sistema de reconocimiento facial puede reducir el tiempo de identificación de un prófugo de semanas a segundos. Un algoritmo de análisis de comunicaciones digitales puede detectar patrones en volúmenes de información que ningún equipo humano podría procesar en tiempo útil. Un sistema de evaluación de riesgo puede aportar un elemento más de análisis al fiscal o al juez al momento de resolver una medida cautelar. Estos beneficios son reales y no deben soslayarse.

Sin embargo, la utilidad de estas herramientas como apoyo a la labor humana no implica que puedan operar de manera autónoma ni que sus resultados puedan sustituir el razonamiento jurídico del operador. El fiscal que solicita una medida cautelar, el juez que la concede y el defensor que la controvierte deben poder comprender y valorar cada elemento del cuadro probatorio. Es precisamente ahí donde la opacidad algorítmica plantea el problema que este trabajo examina: no en el uso de la IA como herramienta, sino en la posibilidad de que sus resultados opacos sean utilizados como fundamento autónomo de decisiones penales con efectos sobre derechos fundamentales.



II. El principio de legalidad y el debido proceso

Cuando un sistema de IA, que opera con opacidad algorítmica, es utilizado por el Estado para investigaciones penales, amplía en forma desmedida su poder punitivo, en desmedro de las libertades individuales. Este trabajo se concentra, particularmente, en la implicancia a nivel de legalidad y debido proceso que la opacidad algorítmica puede traer.

a) El principio de legalidad: taxatividad y previsibilidad

El principio de legalidad penal, consagrado en el artículo 18 de la Constitución Nacional³ y en los artículos 9 de la CADH y 15 del PIDCP⁴, constituye uno de los pilares del Estado de Derecho en materia punitiva. En su formulación clásica, exige que la norma penal sea previa al hecho (*lexpraevia*), escrita (*lexscripta*), estricta (*lexstricta*) y cierta (*lexcerta*). Es esta última exigencia, la de certeza o taxatividad, la que reviste mayor relevancia para el problema que este trabajo aborda.

La taxatividad impone al legislador el deber de describir con suficiente precisión las conductas penalmente relevantes, de modo que el ciudadano pueda anticipar cuáles comportamientos son prohibidos y que el juez pueda verificar si la conducta concreta encuadra en la norma sin acudir a valoraciones arbitrarias.⁵ Roxin señala que la *lexcerta* no es una exigencia puramente técnica de redacción legislativa, sino una garantía material de libertad: la norma penal vaga o indeterminada amplía discrecionalmente el poder punitivo del Estado y priva al ciudadano de la posibilidad de orientar su conducta con arreglo al derecho.

Pero la taxatividad no agota sus efectos en la norma sustantiva. En su dimensión procesal, la exigencia de certeza opera también sobre la *imputación concreta*: el acusado debe poder conocer con precisión de qué se lo acusa, sobre qué base fáctica descansa esa acusación y cuál es el fundamento probatorio de

³Artículo 18, Constitución Nacional Argentina: "*Nadie puede ser penado sin juicio previo fundado en ley anterior al hecho del proceso...*".

⁴Artículo 9 de la Convención Americana sobre Derechos Humanos (CADH): "*Nadie puede ser condenado por acciones u omisiones que en el momento de cometerse no fueran delictivos según el derecho aplicable.*" Artículo 15 del Pacto Internacional de Derechos Civiles y Políticos (PIDCP).

⁵ROXIN, Claus, *Derecho Penal. Parte General*, Tomo I, Civitas, Madrid, 1997, pág. 140 y ss.



cada uno de sus elementos. Una imputación que el propio imputado no puede comprender ni refutar no satisface el estándar constitucional, con independencia de que la norma sustantiva sea precisa.⁶

b) El debido proceso: defensa en juicio, motivación y sana crítica racional

El debido proceso, consagrado en el artículo 18 de la Constitución Nacional y desarrollado en el artículo 8 de la CADH, no es una garantía única sino un conjunto articulado de derechos que aseguran la racionalidad, la contradicción y el control del ejercicio del poder punitivo. Frente al uso de IA en la investigación penal, tres de sus dimensiones resultan especialmente relevantes: el derecho de defensa en juicio, el deber de motivación de las decisiones judiciales y la sana crítica racional como sistema de valoración probatoria.

En cuanto al *derecho de defensa en juicio*, su núcleo material consiste en la posibilidad real —no meramente formal— de conocer la acusación en todos sus términos, acceder a los elementos de prueba de cargo y controvertirlos con plena efectividad.⁷ Este derecho presupone que la prueba sea cognoscible y refutable: la defensa no puede impugnar aquello que no comprende. Cuando una imputación se construye sobre la base de un resultado algorítmico opaco, el derecho de defensa queda vaciado de contenido real. El imputado sabe que el sistema lo señaló, pero no puede saber por qué ni demostrar que el señalamiento es erróneo. Esta asimetría estructural entre la acusación —que tiene acceso al sistema— y la defensa —que no puede auditarlo— no es un problema de capacitación técnica: es una violación de la igualdad de armas que el debido proceso garantiza.

En cuanto a la *sana crítica racional* como sistema de valoración probatoria, su exigencia medular es que la convicción del juez sea el resultado de un razonamiento explicitado, coherente con las reglas de la lógica y la experiencia, y

⁶ZAFFARONI, Eugenio Raúl; ALAGIA, Alejandro; SLOKAR, Alejandro, *Derecho Penal. Parte General*, Ediar, Buenos Aires, 2002, pág. 108 y ss.

⁷Ver MAIER, Julio B., *Derecho Procesal Penal. Fundamentos*, Editorial del Puerto, Buenos Aires, 1999, T. I, pág. 539 y ss.



controlable por las partes y los tribunales superiores. La CSJN estableció en el caso "Casal"⁸ que este sistema impone al tribunal el deber de dar cuenta del camino lógico que lo llevó de la prueba a la conclusión, de manera que ese recorrido pueda ser verificado y eventualmente impugnado. Un resultado algorítmico opaco no puede integrarse válidamente en ese razonamiento porque no es posible dar cuenta de él: no hay camino lógico que explicitar, solo una conclusión sin premisas accesibles. Incorporar ese resultado como elemento de la valoración probatoria equivale a introducir en el razonamiento judicial un eslabón de la cadena que ni el juez ni las partes pueden ver.

Ahora bien, si la defensa no puede auditar la prueba incorporada al proceso y ella no es susceptible de ser explicada mediante las reglas de la sana crítica, inevitablemente llevará a la violación del *deber de motivación de las decisiones judiciales*, descartándolas como una derivación razonada del derecho vigente. en este sentido, la Corte IDH lo ha reconocido el deber de motivación judicial como una garantía autónoma derivada del artículo 8.1 de la CADH: el juez debe expresar las razones de hecho y de derecho que fundan su decisión, de modo que las partes puedan comprenderla y ejercer los recursos correspondientes.⁹ Esta exigencia no se satisface con la mera remisión a un resultado algorítmico. Si el juez escribe en su resolución que concede una medida cautelar porque el sistema de evaluación de riesgo asignó al imputado una puntuación elevada, sin poder explicar por qué esa puntuación es confiable ni en qué datos y razonamientos se funda, la motivación es aparente. El fundamento real de la decisión es la confianza ciega en el algoritmo, que termina por sustituir los argumentos del juez por un proceso algorítmico inaccesible.

III. Opacidad algorítmica: tensiones concretas

⁸Fallos 328:3399, CSJN.

⁹Corte IDH, *Caso Herrera Ulloa vs. Costa Rica*, Excepciones Preliminares, Fondo, Reparaciones y Costas, Sentencia de 2 de julio de 2004, párr. 158.



Los sistemas de IA utilizados en la investigación penal, en particular aquellos basados en técnicas de aprendizaje automático (*machine learning*), operan mediante un proceso que la doctrina ha denominado "caja negra" (*black box*): se conocen los datos que ingresan al sistema y el resultado que produce, pero el proceso de razonamiento interno mediante el cual el sistema arriba a ese resultado no es accesible ni explicable en términos comprensibles para los operadores jurídicos, y en muchos casos tampoco para sus propios desarrolladores.¹⁰

Esta opacidad puede tener distintas causas. En algunos casos es de naturaleza *técnica*: el algoritmo utiliza capas de procesamiento tan complejas que su razonamiento es intrínsecamente irreproducible en lenguaje humano. En otros casos es de naturaleza *comercial*: el código fuente es secreto porque constituye propiedad intelectual de la empresa desarrolladora. En la práctica judicial, ambas formas de opacidad tienden a superponerse, creando una situación en la que ningún actor del proceso —ni el imputado, ni la defensa, ni el fiscal, ni el juez— puede acceder al razonamiento que generó el resultado que se pretende utilizar como evidencia.

La colisión con el principio de legalidad, como ya se vio, opera en tres niveles que es preciso distinguir. En primer lugar, el ciudadano no puede anticipar qué conducta o combinación de conductas resultará penalmente relevante para un sistema de IA. Si un algoritmo de *predictive policing* clasifica a una persona como de alto riesgo delictivo a partir de la combinación de su historial de navegación, sus movimientos georreferenciados, sus patrones de consumo y sus contactos, el ciudadano no puede saber de antemano qué comportamientos cotidianos —todos ellos lícitos individualmente considerados— activarán la atención penal sobre él. Esto es exactamente lo que la taxatividad prohíbe: la indeterminación del presupuesto de la intervención punitiva. En segundo lugar, el imputado no puede comprender ni refutar la imputación, pues no conoce el razonamiento que generó

¹⁰NIEVA FENOLL, Jordi, *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018, pág. 42 y ss.



el resultado algorítmico. No puede señalar el error en la lógica del sistema porque esa lógica es opaca.¹¹En tercer lugar, el juez no puede ejercer el control de legalidad que le impone su función constitucional: cuando los fundamentos de la decisión incluyen un resultado algorítmico que el magistrado no comprende, la decisión se funda, en última instancia, en la confianza ciega en un sistema cuya lógica es inaccesible.

A esos tres niveles de tensión con el principio de legalidad, la opacidad algorítmica agrega tensiones propias con el debido proceso. Desde la perspectiva del derecho de defensa en juicio, la asimetría entre la acusación —que tiene acceso al sistema— y la defensa —que no puede auditarlo— reproduce en el ámbito tecnológico la desigualdad estructural que el proceso acusatorio pretende corregir. Desde la perspectiva de la motivación de las decisiones judiciales, la remisión a un resultado algorítmico no explicado equivale a una motivación aparente: el juez que funda su decisión en el output de una "caja negra" no está motivando, está delegando. Desde la perspectiva de la sana crítica racional, el resultado algorítmico opaco es un eslabón invisible en la cadena del razonamiento probatorio: no puede ser sometido al control lógico que el sistema exige porque carece de premisas accesibles.

Ahora bien, analizando dichas tensiones desde situaciones particulares, se pueden observar algunos ejemplos donde la colisión entre opacidad algorítmica y derechos individuales podrían tener lugar, o bien, ya la tuvieron.

Así, se destaca un caso paradigmático del derecho comparado, el de *State v. Loomis*, resuelto por la Suprema Corte de Wisconsin en 2016.¹² Eric Loomis fue sentenciado a seis años de prisión utilizando, entre otros elementos, la puntuación

¹¹BARONA VILAR, Silvia, *Algoritmización del Derecho y de la Justicia*, Tirant lo Blanch, Valencia, 2021.

¹²Suprema Corte de Wisconsin, *State v. Loomis*, 881 N.W.2d 749 (2016), cert. denied, 137 S. Ct. 2290 (2017). El imputado fue sentenciado a seis años de prisión con base, entre otros elementos, en su alta puntuación de riesgo algorítmico.



de riesgo generada por el sistema COMPAS,¹³ cuyo algoritmo es propiedad de la empresa desarrolladora y no fue revelado a la defensa. La Corte admitió el uso del sistema bajo ciertas condiciones, pero reconoció que el debido proceso podría ser violado en casos futuros si los jueces no comprenden plenamente las limitaciones de la herramienta. La crítica doctrinal señaló con precisión el núcleo del problema: Loomis sabía *qué* información ingresó al sistema, pero no tenía modo alguno de saber *cómo* esa información fue procesada para producir el resultado que fundamentó su condena.

Otro problema que la opacidad presenta es el del fenómeno del *sesgo algorítmico*. Los sistemas de IA entrenados sobre datos históricos de investigaciones penales heredan —y frecuentemente amplifican— los sesgos presentes en esos datos. Si la policía históricamente ha concentrado sus recursos en determinados barrios, grupos étnicos o perfiles socioeconómicos, el algoritmo aprenderá a asociar esas características con la probabilidad delictiva, reproduciendo y escalando la discriminación estructural. Estudios sobre el sistema COMPAS demostraron que el sistema clasificaba erróneamente como de alto riesgo a imputados afroamericanos con una frecuencia casi dos veces mayor que a imputados blancos en situaciones comparables.¹⁴ En América Latina, un informe reciente documentó que un sistema de predicción de riesgo de reincidencia utilizado en Chile arrojó predicciones erróneas en más del 30% de los casos, alentando la detención de personas en razón de sus antecedentes y no de conductas efectivamente cometidas.¹⁵ Este problema tiene una dimensión específica desde la perspectiva de la legalidad penal: la imputación deja de fundarse en la conducta del individuo para fundarse en su pertenencia a un grupo

¹³COMPAS es el acrónimo de *CorrectionalOffender Management ProfilingforAlternativeSanctions*, sistema de evaluación de riesgo de reincidencia desarrollado por la empresa Northpointe Inc., de uso extendido en el sistema de justicia penal de los Estados Unidos.

¹⁴ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren, *"Machine Bias"*, ProPublica, 2016, disponible en: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

¹⁵FairTrials International, *Inteligencia artificial en la seguridad pública y en el sistema penal en América Latina*, 2024, pág. 15.



estadísticamente asociado con la criminalidad, mutando hacia un sistema de justicia penal basado en el derecho penal de autor.

Finalmente, cabe referirse a las *alucinaciones algorítmicas*, fenómeno por el cual los sistemas de IA generan resultados incorrectos con aparente confianza. En los sistemas de reconocimiento facial, esto puede traducirse en la identificación errónea de una persona como prófugo de la justicia, con consecuencias que pueden incluir una detención arbitraria. Un sistema que "ve" semejanza donde no la hay, o que procesa datos históricos incompletos y genera predicciones distorsionadas, no es un instrumento neutral de apoyo investigativo: es una fuente potencial de error sistemático con efectos sobre derechos fundamentales. La capacidad de detectar y corregir estos errores requiere, precisamente, que el sistema sea comprensible y auditable —condición que la opacidad algorítmica niega— y que los datos de entrenamiento sean precisos, representativos y libres de prejuicios.

En síntesis, para evitar estos conflictos, tanto el principio de legalidad como el debido proceso convergen en una misma exigencia frente a la IA: que los resultados algorítmicos utilizados en la investigación y la imputación penal sean explicables, auditables y racionalmente integrables en el razonamiento jurídico. La diferencia entre ambas garantías no es de resultado sino de perspectiva: el principio de legalidad protege al ciudadano frente a la indeterminación del presupuesto de la intervención punitiva; el debido proceso protege al imputado frente a la irracionalidad e incontrolabilidad del proceso mediante el cual esa intervención se materializa. Ambas perspectivas son complementarias e igualmente necesarias.

IV. La experiencia argentina: de la Ciudad de Buenos Aires a la Resolución 710/2024

El problema de la opacidad algorítmica en la investigación penal no es ajeno al derecho argentino. Al contrario, existen antecedentes jurisprudenciales y



normativos de primer nivel que permiten sostener que la discusión ya está planteada en el plano doméstico, aunque no haya sido zanjada.

El caso más relevante en la materia es el proceso judicial iniciado contra el Gobierno de la Ciudad Autónoma de Buenos Aires por la implementación del Sistema de Reconocimiento Facial de Prófugos (SRFP). En 2020, el Observatorio de Derecho Informático Argentino (ODIA) y otras organizaciones iniciaron una acción de amparo cuestionando la constitucionalidad y convencionalidad del sistema, que operaba en tiempo real sobre las imágenes captadas por las cámaras de videovigilancia de la Ciudad y tenía acceso a datos biométricos de aproximadamente diez millones de personas. La presentación señalaba que el sistema operaba con programas que presentaban sesgos discriminatorios por raza, color y etnia, y que se había implementado sin la realización de una evaluación de impacto sobre los derechos de los habitantes.

Una pericia judicial reveló, entre otras irregularidades graves, que el Gobierno de la Ciudad había violado la Ley 5688, que limita el uso del sistema a la búsqueda de personas incluidas en la lista CONARC o en una orden judicial expresa, y que la Policía de la Ciudad había realizado consultas a la base de datos del RENAPER sobre siete millones y medio de personas, excediendo ampliamente ese mandato. Con base en ello, el Juzgado de Primera Instancia en lo Contencioso Administrativo y Tributario N° 4 dictó sentencia el 7 de septiembre de 2022 declarando la inconstitucionalidad de la implementación del SRFP y condicionando su eventual reactivación al cumplimiento de una serie de requisitos, entre los que se destacaban la realización de una evaluación de impacto en protección de datos y la constitución de organismos de control.¹⁶

La Sala I de la Cámara de Apelaciones confirmó esa sentencia el 28 de abril de 2023, rechazando los recursos del Gobierno de la Ciudad y del Ministerio Público Fiscal. La Cámara estableció que el Gobierno de la Ciudad no puede reactivar el sistema hasta tanto: i) se constituyan y funcionen los organismos de

¹⁶O.D.I.A. y otros c/ GCBA s/ Amparo, Juzgado de Primera Instancia en lo Contencioso Administrativo y Tributario N° 4, CABA, 07/09/2022, y Cámara de Apelaciones en lo CATyRC de la Ciudad de Buenos Aires, Sala I, 28/04/2023 (Expte. N° 182908/2020-0).



control; ii) se realicen las investigaciones necesarias para determinar si el sistema tiene impacto diferenciado según características personales de los afectados; iii) se realice una evaluación de impacto en protección de datos; iv) se convoque a la ciudadanía a debatir su implementación.

Lo más relevante para el objeto de este trabajo es que tanto la sentencia de primera instancia como la de Cámara vincularon expresamente el funcionamiento del sistema sin mecanismos de control con una violación al principio de legalidad y a las garantías del debido proceso que deben regir todo accionar de la Administración con incidencia sobre derechos fundamentales. Este precedente es de especial valor porque acredita que el problema de la opacidad algorítmica ya ha sido reconocido por la jurisprudencia argentina como una cuestión constitucional, no meramente de privacidad o protección de datos.

En el plano normativo federal, la Resolución 710/2024 del Ministerio de Seguridad de la Nación creó la UIAAS con la misión de prevenir, detectar, investigar y perseguir el delito mediante el uso de IA.¹⁷ Entre sus funciones se incluyen el patrullaje de redes sociales y la Dark Web, la identificación y comparación de imágenes, el análisis de cámaras de seguridad en tiempo real mediante reconocimiento facial, y la utilización de algoritmos de aprendizaje automático para predecir delitos futuros a partir de datos históricos. La norma no contiene previsión alguna sobre la explicabilidad de los sistemas utilizados, la auditabilidad de sus resultados, la supervisión judicial específica de la actividad algorítmica ni el derecho de la defensa a conocer el funcionamiento de los sistemas empleados en la investigación. Este vacío compromete simultáneamente las exigencias del principio de legalidad y las garantías del debido proceso, y no es un detalle técnico: es el núcleo del problema constitucional que este trabajo identifica.

V. Estándares aplicables y criterios mínimos

¹⁷Ministerio de Seguridad de la Nación, Resolución 710/2024, publicada en el Boletín Oficial de la República Argentina el 29/07/2024.



Establecido el problema, corresponde ahora identificar los estándares jurídicos aplicables al derecho argentino y los criterios mínimos que, derivados de esos estándares, deben satisfacer los sistemas de IA utilizados en la investigación penal.

En el plano convencional, la jurisprudencia de la Corte IDH provee el piso mínimo exigible tanto desde el principio de legalidad como desde el debido proceso. En materia de legalidad, la Corte ha sostenido de manera reiterada que el Estado debe precisar normativamente las conductas relevantes y fundar sus decisiones en circunstancias objetivas y ciertas del caso concreto.¹⁸ En materia de debido proceso, ha establecido que la motivación de las decisiones judiciales es una garantía autónoma del artículo 8.1 de la CADH¹⁹ y que el acusado debe tener la posibilidad efectiva de ejercer el derecho de contradicción respecto de cada elemento de cargo. Ninguno de estos estándares puede satisfacerse cuando el fundamento de la imputación es un resultado algorítmico opaco que ninguno de los actores del proceso puede comprender ni controlar.

Desde el sistema universal, el Alto Comisionado de las Naciones Unidas para los Derechos Humanos ha sostenido que las iniciativas regulatorias deben exigir una evaluación de los riesgos e impactos para los derechos humanos que conlleven los sistemas de IA antes, durante y después de su uso, y que deben existir garantías de transparencia, supervisión independiente y acceso a remedios efectivos, de manera especial cuando es el propio Estado el que utiliza estas tecnologías. Las tecnologías de IA que no puedan ser operadas en cumplimiento con la normativa internacional de derechos humanos deben ser prohibidas o suspendidas hasta que se implementen protecciones adecuadas.²⁰

¹⁸Corte IDH, *Caso Baena Ricardo y otros vs. Panamá*, Sentencia de 2 de febrero de 2001, Serie C N° 72, párr. 107.

¹⁹Corte IDH, *Caso Chaparro Álvarez y Lapo Íñiguez vs. Ecuador*, Sentencia de 21 de noviembre de 2007, Serie C N° 170, párr. 93.

²⁰Alto Comisionado de las Naciones Unidas para los Derechos Humanos, *"La inteligencia artificial debe tomar como base los derechos humanos"*, julio de 2023, disponible en: <https://www.ohchr.org>.



En el plano del derecho comparado, el *AI Act* europeo clasifica como sistemas de alto riesgo a todos los utilizados en la aplicación de la ley y la administración de justicia, y exige para ellos: supervisión humana obligatoria, documentación técnica completa y accesible, auditorías independientes, registro de decisiones y evaluación previa de sesgos. Establece además que las autoridades policiales no podrán tomar ninguna decisión con efectos jurídicos adversos sobre una persona basándose *únicamente* en resultados de sistemas de identificación biométrica a distancia.²¹ Aunque este reglamento no tiene aplicación directa en Argentina, sus estándares son coherentes con los que pueden derivarse del propio bloque de constitucionalidad argentino y ofrecen un parámetro técnico-jurídico valioso para la interpretación.

De la articulación de estos estándares pueden derivarse los siguientes criterios mínimos que deben satisfacer los sistemas de IA en la investigación penal:

i) Explicabilidad obligatoria. Ningún resultado de un sistema de IA puede servir como fundamento —total o parcial— de una imputación, medida cautelar o elemento de cargo si no puede ser explicado en términos comprensibles por el órgano jurisdiccional. Este criterio deriva simultáneamente del principio de legalidad —en su exigencia de certeza sobre la base de la imputación— y del debido proceso —en su exigencia de motivación real y de valoración racional de la prueba.

ii) Auditabilidad independiente. La defensa debe tener acceso a una pericia independiente sobre el funcionamiento del sistema, incluyendo documentación técnica suficiente para evaluar su lógica, su tasa de error y la presencia de sesgos. Este criterio es una exigencia directa del derecho de defensa en juicio: el secreto comercial no puede oponerse válidamente frente a las garantías constitucionales del imputado.

²¹Reglamento (UE) 2024/1689, artículo 26, apartado 10. La disposición establece que las autoridades policiales no podrán tomar ninguna decisión que produzca un efecto jurídico adverso sobre una persona basándose *únicamente* en los resultados de sistemas de identificación biométrica a distancia.



iii) Prohibición de decisiones adversas fundadas exclusivamente en IA.

Ninguna decisión con efectos jurídicos adversos sobre el imputado puede fundarse de manera exclusiva en un resultado algorítmico. Los resultados de IA pueden constituir un elemento más dentro de un cuadro probatorio más amplio, pero nunca el único ni el determinante fundamento de una medida que restrinja derechos fundamentales.

iv) Evaluación de impacto previa. Antes de incorporar cualquier sistema de IA a la investigación penal, el Estado debe realizar una evaluación de impacto que incluya análisis de sesgos, tasa de error verificada empíricamente, impacto diferenciado por grupos y compatibilidad con los derechos fundamentales. Este es el criterio consagrado en la sentencia del caso ODIA c/ GCBA.

v) Control judicial efectivo. El juez que interviene en el proceso donde se utiliza IA debe poder ejercer un control real —no meramente formal— sobre la legalidad y racionalidad de la herramienta empleada. Ello requiere que el sistema sea explicable o que existan mecanismos de supervisión independiente que permitan al tribunal formarse un juicio fundado sobre la confiabilidad del resultado.

Es preciso señalar que estos criterios no requieren de una nueva legislación para ser operativos en el derecho argentino, ya que, derivan directamente del principio de legalidad, debido proceso y, en general, del bloque de constitucionalidad federal. Sin embargo, a fin de evitar interpretaciones judiciales arbitrarias y/o contradictorias, su inclusión en una norma que regule de una manera concentrada y particular dichos estándares, traería grandes beneficios.

Conclusión

Intentar precisar los límites que el ordenamiento jurídico impone al uso de la inteligencia artificial en la investigación penal es una tarea que apenas comienza y que debe ser motivo de una ardua labor parlamentaria y judicial en nuestro país. Los avances tecnológicos se presentan con una velocidad que el derecho difícilmente puede acompañar de manera inmediata, y los operadores jurídicos se ven frecuentemente ante situaciones para las cuales las normas existentes no



ofrecen respuesta explícita. El presente trabajo no pretende agotar un debate que, en el plano internacional, acumula ya una doctrina considerable y que en el plano argentino se encuentra, en sus aspectos más específicos, en sus primeras etapas.

Sin perjuicio de ello, como consecuencia del desarrollo efectuado, pueden extraerse las siguientes conclusiones.

En primer lugar, el uso de sistemas de IA como herramienta de apoyo a la investigación penal no es en sí mismo incompatible con las garantías constitucionales; de hecho, resultaría totalmente positivo. Países de la región como Brasil y Colombia, han incorporado el uso de sistemas de IA al servicio de justicia con buenos resultados. La IA puede aportar beneficios concretos en términos de eficiencia y precisión investigativa, siempre que sea concebida como un instrumento al servicio del razonamiento humano y no como un sustituto de él. El operador jurídico —fiscal, juez, defensor— debe conservar en todo momento el control racional y jurídico del proceso.

En segundo lugar, la opacidad algorítmica no es un déficit técnico superable mediante peritaje: es una incompatibilidad estructural con las exigencias del Estado de Derecho en materia penal. Cuando el razonamiento que sustenta un resultado de IA no puede ser explicado, auditado ni verificado, ese resultado no puede servir como fundamento de una imputación que satisfaga el estándar de taxatividad exigido por el principio de legalidad ni las exigencias de motivación racional y contradicción efectiva que impone el debido proceso.

En tercer lugar, el fenómeno de las alucinaciones algorítmicas y el sesgo discriminatorio derivado del entrenamiento con datos históricos son riesgos concretos y documentados que no pueden ser ignorados. Ambos comprometen no solo la exactitud de los resultados sino su validez constitucional, en la medida en que generan imputaciones fundadas en características estadísticas de grupos antes que en conductas individuales, vulnerando el principio de culpabilidad por el hecho y constituyéndose como herramienta de potencial aumento de la discriminación. Para evitar sesgos involuntarios o inexactitudes, es clave garantizar que los conjuntos de datos utilizados por los sistemas de IA sean



precisos, representativos y libres de prejuicios. Para ello, se ha propuesto²² que los sistemas de IA sean diseñados bajo un enfoque de "caja blanca", que garantice que las decisiones sean razonadas y fundadas y que permita determinar con precisión el paso a paso del proceso de toma de decisiones. La inteligibilidad y trazabilidad de estos procesos deben estar plenamente documentadas.

En cuarto lugar, el derecho argentino ya dispone de los estándares necesarios para dar una respuesta a este problema, sin necesidad de nueva legislación. El principio de legalidad, debido proceso y el bloque de constitucionalidad federal en general permiten derivar los criterios de explicabilidad, auditabilidad, prohibición de decisiones exclusivamente algorítmicas, evaluación previa de impacto y control judicial efectivo como condiciones de validez constitucional del uso de IA en la investigación penal, aunque su regulación en una ley que los concentre, traería beneficios a la ciudadanía y limitaría la amplia discrecionalidad judicial en la materia.

En síntesis, ni el principio de legalidad ni el debido proceso prohíben el uso de inteligencia artificial en la investigación penal. Lo que prohíben es que sus resultados opacos, sesgados o inverificables sean utilizados como fundamento de la persecución penal sin que el imputado, su defensa y el juez puedan comprenderlos, cuestionarlos y controlarlos. Esa es la frontera que el Estado constitucional de Derecho impone, y es el desafío que los operadores jurídicos argentinos tienen ante sí.

Bibliografía consultada.

1. ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren, "Machine Bias", ProPublica, 2016.
2. BARONA VILAR, Silvia, *Algoritmización del Derecho y de la Justicia*, Tirant lo Blanch, Valencia, 2021.

²²Cristallo, J.; Lavin, R.; Gayraud, F. y Daglio, A. C. (2023). *Transformación digital en el sistema de justicia: oportunidades para una justicia rápida, accesible y transparente*. Buenos Aires: Fundar. Disponible en <https://www.fund.ar>.



3. Cristallo, J.; Lavin, R.; Gayraud, F. y Daglio, A. C., *Transformación digital en el sistema de justicia*, Fundar, Buenos Aires, 2023.
4. DE LUIS GARCÍA, Elena, "Justicia, inteligencia artificial y derecho de defensa", IDP. Revista de Internet, Derecho y Política, n° 39, 2023.
5. FairTrials International, *Inteligencia artificial en la seguridad pública y en el sistema penal en América Latina*, 2024.
6. FERRAJOLI, Luigi, *Derecho y Razón. Teoría del Garantismo Penal*, Trotta, Madrid, 1998.
7. Harvard Journal of Law&Technology, "AlgorithmicDueProcess: MistakenAccountability and Attribution in State v. Loomis", 2017.
8. MAIER, Julio B., *Derecho Procesal Penal. Fundamentos*, Editorial del Puerto, Buenos Aires, 1999, T. I.
9. NIEVA FENOLL, Jordi, *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018.
10. ROXIN, Claus, *Derecho Penal. Parte General*, Tomo I, Civitas, Madrid, 1997.
11. ZAFFARONI, Eugenio Raúl; ALAGIA, Alejandro; SLOKAR, Alejandro, *Derecho Penal. Parte General*, Ediar, Buenos Aires, 2002.

Jurisprudencia citada.

1. CSJN, "Fallos" 328:3399 ("Casal", 2005).
2. Corte IDH, "Herrera Ulloa vs. Costa Rica", Sentencia de 02/07/2004.
3. Corte IDH, "Chaparro Álvarez y Lapo Íñiguez vs. Ecuador", Sentencia de 21/11/2007.
4. Corte IDH, "Baena Ricardo y otros vs. Panamá", Sentencia de 02/02/2001.
5. ODIA y otros c/ GCBA s/ Amparo, Juzgado CAyT N° 4, CABA, 07/09/2022.
6. Cámara de Apelaciones CATyRC de la CABA, Sala I, 28/04/2023 (Expte. N° 182908/2020-0).
7. Suprema Corte de Wisconsin, Estados Unidos, *State v. Loomis*, 881 N.W.2d 749 (2016).