

“XXXII CONGRESO NACIONAL DE DERECHO PROCESAL”

“El Derecho Procesal a 30 años de la reforma a la Constitución Nacional”.

Mar del Plata, 14 al 16 de Noviembre de 2024

Comisión de derecho procesal penal:

1- Inteligencia artificial: proceso penal y derechos fundamentales (arts. 75, inc. 19 y 22, CN)

Ponencia General:

“Justicia Algorítmica y Prisión Provisional: Dilemas y Límites del uso de la inteligencia artificial en las valoraciones judiciales de riesgo procesal”

Presentada por

Gabriel Adrián Bombini¹

¹ Prof. e Investigador en Derecho Penal y Criminología de la Facultad de Derecho de la Universidad Nacional de Mar del Plata. Juez de Garantías N° 5, Departamento Judicial Mar del Plata.

1. Problematicaciones: ¿Son admisibles y eficaces los programas de inteligencia artificial para la valoración judicial del riesgo procesal en los procedimientos penales?

La extraordinaria difusión de una diversidad de herramientas de inteligencia artificial ha impregnado de un modo decisivo prácticamente todos los aspectos de la vida social a esta altura del Siglo XXI.

Frente a las múltiples ventajas que ofrecen estos mecanismos en términos de disponibilidad y celeridad en el uso de la información y resolución de plurales problemas de la vida cotidiana y de diversas actividades sociales, se erigen un conjunto de dilemas éticos que desafían a nuestras sociedades en torno a los límites y las derivas del desarrollo tecnológico².

A este escenario polémico no resultan ajenos los debates en el campo del derecho o de las disciplinas jurídicas, ni aún menos, las propias instituciones de la administración de justicia³. En diversos contextos, se van conociendo y acumulando un número cada vez más significativo de experiencias y prácticas que, con menor o mayor nivel de institucionalización, apelan a la utilización de herramientas de Inteligencia Artificial, en la administración de justicia en general, y en el sector de la Justicia Penal, en particular.

La inusitada vertiginosidad con las que se despliegan las transformaciones del presente, pretenden presentar a la administración de justicia del siglo XXI como *innovativa, celera, productiva, eficiente y transparente a partir de la incorporación de esas cualidades propias de una sociedad digital sustentada en la tecnología*⁴, pero rezagan dos dimensiones

² Por todos, HAN, Byung-Chul *No-cosas. Quiebras del mundo de hoy*. Colonia: Penguin-Taurus, 2021; HAN, Byung-Chul, *En el enjambre*, Barcelona, Herder, 2014, MOROZOV. E., *La locura del solucionismo tecnológico*, Madrid, Katz Editores, 2015.

³ Al respecto, cfr. por todos trabajos como los de NIEVA FENOLL, Jordi, *Inteligencia artificial y proceso judicial*. Madrid: Marcial Pons, 2018 o de ALVAREZ LARRONDO, Federico (Dir.) *Inteligencia artificial y derecho*. Buenos Aires, Hammurabi, 2020; o en particular para el ámbito penal de MIRÓ LINARES, Fernando "Inteligencia artificial y Justicia Penal: Más allá de los resultados lesivos causados por robots" en *Revista de Derecho Penal y Criminología*, n. 20, pp. 87-130, 2018; RIQUERT, Marcelo. *Inteligencia artificial y Derecho Penal*. Buenos Aires: Ediar, 2022.

⁴ BARONA VILAR, Silvia, *Algoritmización del Derecho y de la justicia. De la Inteligencia Artificial a la Smart Justice*, Valencia, Tirant lo Blanch, 2021.

cruciales en las discusiones sobre el uso de la tecnología en estos complejos espacios institucionales.

Por un lado, las necesarias reflexiones éticas y filosóficas en torno a los límites y reparos que deben imponerse en el campo de la actividad jurisdiccional en tanto que las decisiones judiciales dirimen aspectos ligados a valores constitucionales, antropológicos y políticos que son inherentes a esa función.

Desde otro enfoque, vinculado a una dimensión del ser, no puede dejar de advertirse que el encantamiento que produce la incorporación de novedosos y sofisticados instrumentos tecnológicos, con frecuencia suele relegar las preguntas utilitaristas sobre su real eficacia en términos de aporte de elementos que permitan proveer respuestas judiciales más certeras, uniformes y ajustadas a la realidad de los casos concretos que se gestionan.

Estos interrogantes, se potencian en la campo de la Justicia Penal, cuando el uso de la Inteligencia Artificial se explora en relación a aquellas decisiones judiciales que impactan sustancialmente en las personas sometidas a un proceso penal en tanto se proyectan sobre su libertad ambulatoria, y por ende, resultan susceptibles de afectar derechos fundamentales y garantías constitucionales.

Desde esta perspectiva, tanto a la hora de la determinación judicial de la pena a imponer a las personas condenas (*sentencing*) como a la de dirimir las medidas cautelares necesarias para garantizar los fines procesales, y en particular la necesidad de la prisión provisional, en distintas jurisdicciones se han venido incorporando programas de IA que despiertan resquemores y prevenciones de todo tipo.

En esta oportunidad, esta ponencia, se centrará en el segundo de los problemas aludidos, asumiendo ambas dimensiones anunciadas; esto es: a) si en las valoraciones judiciales de riesgo procesal al momento de adoptar medidas cautelares en el curso del proceso penal -en particular la posibilidad de la prisión provisional o preventiva- es admisible acudir al uso de programas o herramientas de IA, y b) si en esa labor tales mecanismos resultan eficaces en términos de la esperada objetividad, certeza y reducción del margen de error que debería brindar la sofisticación de los recursos tecnológicos.

Con esos objetivos, el trabajo se estructurará en los tres apartados subsiguientes. En el primero, se reflexionará en forma breve sobre la emergencia de lo que oportunamente se denominó un modelo de *justicia actuarial*, y en particular sobre la creación de espacios institucionales dedicados a la provisión de información sobre cálculos de riesgo a jueces y juezas para la toma de decisiones sobre las medidas cautelares. Allí se reseñarán muy escuetamente aspectos vinculados a las legitimaciones de estas experiencias, como así también a su funcionamiento y efectos prácticos.

En el siguiente apartado se intentarán trazar -también con la brevedad del caso- algunas pinceladas del surgimiento de lo que se ha dado en llamar una *justicia algorítmica* y en particular su utilización en este campo de la prisión cautelar.

Finalmente, se dejarán asentadas algunas posiciones provisionales sobre los límites ético-constitucionales y asimismo, los de tipo utilitarista que emergen a partir de la pretensión del uso de la IA en el proceso penal, y en especial en relación a la prisión preventiva.

2. Justicia Actuarial: orígenes y dilemas de la apelación a mediciones y cálculos de riesgo en la valoración de la peligrosidad procesal

Desde hace tiempo que se ha planteado la idea que la justicia criminal ha reconvertido sus objetivos sociales de transformación individual para la prevención del delito hacia otros de tipo institucional vinculados a la gestión y gerenciamiento eficaz de casos y contención de grupos poblacionales riesgos.

En esa línea, a mediados de la década de los años 90 del siglo pasado, Malcom Feeley y Jonathan Simon sugirieron la emergencia de una nueva penología que impregna las lógicas del sistema de justicia penal de carácter *actuarialista*. De acuerdo con estos autores, los sistemas penales comenzaron a reorientarse en función de nuevos discursos basados en el lenguaje de la probabilidad y el riesgo, desplazando las viejas categorías del diagnóstico clínico y el tratamiento resocializador.

La misión del derecho penal relegaría aquella de la tradición liberal destinada a proteger al individuo frente al estado, para favorecer una

adecuada clasificación de categorías de sujetos peligrosos y de ese modo regular el delito a límites funcionalmente tolerables.⁵

De acuerdo con esta descripción, la *justicia actuarial* abandona el viejo lenguaje médico tratamental, para guiarse por una *racionalidad administrativa* que, readaptando las técnicas actuariales del cálculo y predicción del riesgo propias del mundo de los seguros y el derecho de daños, desarrolla lógicas y prácticas de gerenciamiento empresarial (“management”), persiguiendo una gestión económicamente eficiente de grupos poblacionales de riesgo a través de medidas penales de control en detención o en el medio social⁶.

El propio Feeley rastrea en los orígenes de la justicia actuarial⁷, vinculándolo con sus investigaciones sobre la centralidad de las decisiones iniciales del proceso penal (*“modelo de proceso previo al juicio” -pretrial process model-*) y sus análisis de los proyectos liberales de aquella época en los Estados Unidos de Norteamérica, de creación de oficinas de asistencia a los tribunales para la evaluación de alternativas al encarcelamiento preventivo⁸.

En particular, reseña diversas experiencias locales que culminaron por la implementación en el nivel federal de “Agencias de Servicios Previos al

⁵ SIMON, Jonathan – FEELEY, Malcom, “La nueva penología: notas acerca de las estrategias emergentes en el sistema penal y sus implicaciones”, en Revista Delito y Sociedad, No 7, 1995, original en inglés “The new penology: notes on the emerging strategy of corrections and its implications”, *Criminology*, 1992.

Véase sobre el debate de la justicia actuarial y el cálculo de riesgo, entre muchas otras, las contribuciones de HALLSWORTH, Simon, “Repensando el Giro Punitivo”. En Revista Delito y Sociedad Nº 22, año 16, UNL, 2007; O’MALLEY, Pat, “Riesgo, Poder y Prevención del Delito”, en Revista Delito y Sociedad Nº 20, año 13, UNL, 2004; o SPARKS, Richard, “Perspectivas sobre Riesgo y Política Penal” en Revista Delito y Sociedad Nº 23, año 16, UNL, 2007.

⁶ Se utiliza la expresión difundida por FEELEY Y SIMON en los años 90, a partir del mencionado artículo decisivo en el debate de la sociología del castigo. Allí se sugería la emergencia de una nueva lógica penal de carácter selectivo, basada en criterios actuariales y desplegada sobre poblaciones estimadas peligrosas, y que resultaba opuesta a su antecesora “vieja penología”, orientada por la noción de resocialización.

⁷ La línea analítica crítica de debate sobre una nueva penología actuarial o managerial iniciada por FEELEY-SIMON fue retomada por diversos autores como, por ejemplo, BROWNLEE, Ian, “New Labour: New Penology? Punitive Rhetoric and the Limits of Managerialism in Criminal Justice Policy”, *Journal of Law and Society*, Vol. 25, 3/9/1998; BRANDARIZ GARCÍA, José Ángel, “La difusión de las lógicas actuariales y gerenciales en las políticas punitivas”, *InDret* (2), 2014; O LAMAS LEITE, André, “Nueva penología», punitive turn y Derecho Penal: *quo vadimus?* Por los caminos de la incertidumbre (pos)moderna”. *InDret* (2). Ambos disponibles en www.indret.com.

⁸ FEELEY, Malcom, “Reflexiones sobre los orígenes de la justicia actuarial”, en Revista Delito y Sociedad UNL, 2016.

Juicio”, cuyo objetivo principal es asistir al tribunal en “administrar” la liberación previa a través de una orientación predictiva, suministrar información social sobre el arraigo y compromiso familiar o comunitario como modo de reformar el desigual sistema de libertad bajo fianza económica.

Sin embargo, advierte Feeley que rápidamente, se produjo un rápido y giro en las metas de esas reformas de la libertad bajo fianza (“Bail reform”), en tanto que en unos pocos años, *las políticas saltaron de maximizar la libertad antes del juicio a las draconianas provisiones para la detención preventiva*. Así las cosas, estas mismas agencias que habían sido pergeñadas para la predicción del riesgo procesal de fuga o falta de acatamiento a los llamados judiciales futuros, pasaron a encargarse de predecir el *riesgo para la seguridad pública*, validando su tecnología y profesionalismo.

En palabras del autor:

*“...En poco tiempo la misión de las agencias de servicios prejudiciales se transformó de predecir la comparecencia a juicio a predecir la peligrosidad con propósitos de detención preventiva, y de recordar a los liberados de su comparendo pendiente en el tribunal a encargarse de la vigilancia de los liberados de menor riesgo...”*⁹

Es decir que:

*“...una vez que la institución se ha establecido para predecir la comparecencia a juicio, sólo ha sido un pequeño paso pedirle que prediga la peligrosidad... reconozcamos que esta institución no opera en el sentido “fuerte”, no emplea precisos y ajustados análisis estadísticos. Pero ... La pátina de ciencia, tecnología y predicción ha legitimado una práctica que históricamente ha sido extraña a la tradición judicial angloamericana. Ha llevado a nuevos objetivos, nuevas técnicas y nuevas instituciones. Ha tenido consecuencias reales y extensas...”*¹⁰

Este antecedente relevante de transformación en la justicia penal reciente, se potencia y recrea a partir del fuerte despliegue y desarrollo de la tecnología y los modelos de inteligencia artificial en los años sucesivos, dando lugar a reflexiones sobre la consagración de una *justicia algorítmica*, unas de cuyas derivaciones relevantes, resulta -en la misma dirección de la justicia

⁹ 2016, Ob.cit.

¹⁰ Idem.

actuarial- en la medición y predicción de riesgos procesales, claro que con una *pátina* más sofisticada. Veamos.

3. Justicia Algorítmica: experiencias de uso de la inteligencia artificial en la valoración del riesgo procesal para el dictado de medidas cautelares y la prisión provisional

En efecto, en el tiempo reciente, un conjunto de avances tecnológicos de enorme relieve ha permitido consagrar la expresión Inteligencia Artificial (IA), reemplazando la actividad humana y abarcando *“cualquier sistema que muestre un comportamiento inteligente mediante el análisis de su entorno y la adopción de medidas –con cierto grado de autonomía– para lograr objetivos específicos”*.¹¹

A su vez, dentro del espectro más amplio de los mecanismos de IA aplicables al campo del derecho y la administración de justicia, emergen una subcategoría específica denominada *“Sistemas de Inteligencia Artificial Judicial para la Valoración de Riesgo”*, también llamados en el ámbito anglosajón como *Pretrial Risk Assessment Tool (PRAT)*.¹²

De acuerdo con Serventich, este sistema, que replica la mente humana, es capaz de procesar e interpretar una vasta base de datos de información general (relativa a causas similares, fallos anteriores y datos estadísticos) y particular (vinculada al investigado, al delito y a las circunstancias del caso *sub iudice*) a la luz de lo que nuestra normativa, doctrina y jurisprudencia han entendido como indicadores objetivos de riesgos procesales¹³, y con ello

¹¹ SERVENTICH Catalina, “Inteligencia artificial en el proceso penal. ¿Más vale humano conocido o algoritmo por conocer?” en *Revista Jurídica Austral* Vol. 3, N° 2, Diciembre, 2022: 869-880: con cita de Miro Linares, 2022, pág. 2.

¹² Conf. SERVENTICH ob. cit.; con referencia a SIMÓN CASTELLANO, Pere. *Justicia cautelar e inteligencia artificial*. La alternativa a los atávicos heurísticos judiciales, Barcelona: Bosch, 2021, p. 130.

¹³ Siguiendo a la autora, podrían mencionarse: *“...(i) respecto al entorpecimiento de la investigación, en la peligrosidad de la conducta delictiva en su conjunto, en la posición de poder que podría ejercer en una estructura organizativa, en la posibilidad de colaboración de otros en la destrucción de pruebas o en el carácter embrionario de la investigación; (ii) respecto al riesgo de fuga, en el arraigo familiar, en la existencia de fugas anteriores, en la pertenencia a una banda organizada que podría colaborar en su fuga, en los medios económicos de los que disponga o en su voluntad de permanecer localizable durante el proceso...”* (SERVENTICH, ob. cit. 872).

puede brindar un pronóstico de mayor fiabilidad (libre de afectaciones) en la materia en cuestión.

Ragués i Vallés destaca que en estados como Kentucky o California desde hace algunos años se emplean para la decisión judicial de arresto cautelar, las denominadas *risk assessment tools o bail algorithms*.

Estas herramientas resultan una suerte de fórmulas matemáticas predictivas basadas en la experiencia de casos anteriores que, tomando en cuenta determinados datos del caso concreto y del sujeto sospechoso, tales como el delito presuntamente cometido, las circunstancias familiares o profesionales, antecedentes, posibles adicciones, efectúan juicios de probabilidad de huida o reincidencia¹⁴.

Sus promotores destacan diversos aspectos que tornarían favorable su implementación, los cuales no están exentos de polémica. Señala Ragués i Vallés¹⁵, por un lado, a la posibilidad que algunas investigaciones reflejan en relación a la expectativa que utilización impacte positivamente en la reducción del uso cuantitativo de la prisión preventiva, objetivo que viene controvertido por estudios que indican el efecto contrario¹⁶. Por otro, el carácter más científico y libre de prejuicios de la decisión judicial, así como las mayores garantías que supone si se lo compara con el tradicional sistema de fianzas, muy desfavorable para los menos privilegiados económicamente.

Sin embargo, la virtuosidad de la incorporación de los sistemas de IA estaría dada no sólo por la cantidad de información disponible y procesable ampliamente superior a las capacidades humanas, sino también por la superación de los déficits que tradicionalmente se le adjudican a decisiones adoptadas de forma rápida, con escasa información y actividad probatoria nula o muy limitada en los momentos iniciales del proceso judicial, tal como ocurre con las medidas cautelares¹⁷.

¹⁴ Cfr. RAGUES I VALLES, Ramon, *La prisión provisional como ultima ratio*. Madrid. Marcial Pons, 2023: pág. 216-217. Destaca el autor que para el año 2020 el uso de estas herramientas, empleadas también para valorar riesgos en casos en fase de ejecución penal, alcanzaba ya a un 25% de la población penitenciaria estadounidense.

¹⁵ Ob. cit. pág. 218.

¹⁶ Idem, pág. 218, nota. 100.

¹⁷ NEIRA PENA, Ana María "Inteligencia artificial y tutela cautelar. Especial referencia a la prisión provisional" en *Rev. Bras. de Direito Processual Penal*, Porto Alegre, v. 7, n. 3, p. 1897-1933, set.-dez. 2021.

A su vez, la aspiración de objetividad del abordaje artificial, despojaría a la decisión de variables emocionales, condicionadas por aspectos como la apariencia o el lenguaje empleado por el encausado, o por estandarización respecto de casos similares pasados¹⁸.

Explica Neira Pena, que las decisiones cautelares son adoptadas, muchas veces, de forma casi automática, atendiendo a la experiencia reciente del juez, que otorgará una especial relevancia, en ocasiones desproporcionada, a determinadas variables –como podrían ser la nacionalidad del encausado y/o la dinámica delictiva observada- por el mero hecho de estar presentes en casos decididos recientemente, o juzgados en el pasado y valorados como representativos, sin contrapesar adecuadamente otras variables distintas y particulares del nuevo caso a enjuiciar (ob. cit. 1905).

En este sentido, se postula que la presunta ventaja que podría aportar la IA, en tanto que método automatizado de valoración de riesgos basado en evidencias y datos objetivos -y no en las meras intuiciones y apreciaciones personales- es, precisamente, el avance hacia la superación de los sesgos existentes en los juzgadores¹⁹.

Incluso, muy particularmente, con relación a las medidas cautelares como la prisión preventiva, se sostiene que la IA puede realizar aportaciones de utilidad en lo relativo a uno de sus presupuestos procesales fundamentales, cuál es el *peligro en la demora* como riesgos concretos que las medidas cautelares tratan de precaver. Aunque su uso en este ámbito en absoluto está exento de cuestionamientos, ni debe de estar desprovisto de prevenciones y cautelas, quienes promueven su difusión mayor en la

¹⁸ NEIRA PENA cita a SIMON CASTELLANO, cuando que expresa que en relación con las decisiones cautelares, “*se produce una suerte de automatismo en relación con casos parecidos o análogos que han sido resueltos con anterioridad, porque ante la falta de actividad probatoria difícilmente se pueden apoyar en algo más. Se refuerza el anclaje basado en la comodidad y la necesaria reducción de la carga de trabajo, la accesibilidad y la disponibilidad en relación con aquellos argumentos que el juzgador tiene más cerca, esto es, los de quien solicita o interesa la medida, es decir, el Ministerio Fiscal, y la representatividad*”. (ob. cit. 1904-1905)

¹⁹ Es lo que, según la autora, BORGES BLÁZQUEZ, Raquel, denomina la “trampa de la neutralidad”, basada en la creencia de que la IA se alimenta única y exclusivamente de cifras y que los números sólo pueden ser tratados de forma matemática, alimentando con esto el mito de su matematización).

administración de justicia penal sostienen que la *“potencialidad de la IA para fundamentar estas valoraciones resulta indudable”*²⁰.

Sin embargo, no debe olvidarse que, en definitiva, tanto las bases de datos que las nutren como los algoritmos en sí mismos, no dejan de ser creaciones humanas y, como tales, los modelos matemáticos en que se sustentan, trasuntan o reflejan objetivos e ideologías²¹.

Es que resultando cierto que la IA permite gestionar de forma eficaz ingentes cantidades de datos, inabarcables para la memoria y el raciocinio humano, ocurre que la forma en que se construye el algoritmo y los datos con que este se entrena responden a decisiones humanas, lo que en ocasiones provoca que los sesgos propios del razonamiento humano sean replicados, e incluso exacerbados, en los resultados del algoritmo²².

Es por eso que la asunción de una óptica prudente y realista en torno a su utilización, no puede pasar por alto que los IA predictivos no están exentos de sesgos y de riesgos²³.

Así destaca Neira Pena que en primer lugar, que el algoritmo o los datos con que el mismo ha sido creado, al replicarlos para su aplicación a nuevos casos, pueden tender a introducir o perpetuar las discriminaciones presentes en la actuación policial y judicial.

Y señala en segundo orden, que su fiabilidad ha demostrado no ser mucho mayor que la de la intuición humana, aunque, en este caso, revestida de un cientificismo, lo que dota a las decisiones de una fuerza categórica cuanto menos preocupante²⁴.

²⁰ NEIRA PENA. Ob. cit.: pág. 1909.

²¹ Cfr. NEIRA PENA, ob. cit. pág. 1902.

²² En el mismo sentido, se pronuncia, MARTÍNEZ GARAY, Lucía (en Peligrosidad, Algoritmos y Due Process: el Caso *State V Loomis*. *Revista de Derecho Penal y Criminología*, 3.a Época, n 20 pp. 485-502, p. 496, julio de 2018) en tanto señala que *“cuando se afirma que la información que proporcionan las herramientas de valoración de riesgos es objetiva, y por ello preferible a la valoración personal del juez, que puede estar sesgada por prejuicios inconscientes (raciales, de clase, de género, etc.), se olvida que los algoritmos son un mero reflejo de los datos con los que se los alimenta, de modo que si esos datos incorporan sesgos el algoritmo los reproducirá, o peor aún, los exacerbará”*. (NEIRA PENA, ob cit. Pág. 1905, y nota 21).

²³ HANNAH-MOFFAT, KELLY – AVILA, FERNANDO, “Prospective but disconnected partners: AI-informed criminal risk prediction” en LINDGREN, SIMON (ed.), *Handbook of Critical Studies of Artificial Intelligence*, Edward Elgar Publishing, 2023.

²⁴ Cfr. Ob. cit. Pág. 1915.

Riquert comenta que desde el lado crítico se enfatiza en que bajo la etiqueta de la “valoración de riesgos” se esconde una versión tecnológica de los pronósticos de “peligrosidad”, que ahora se formulan bajo una predicción actuarial centrada en valores de conjuntos de sujetos con factores análogos, lo que ofrecería un método que, aunque con mayor grado de fiabilidad, no estaría exento de “falsos positivos” (pronósticos de delincuencia no concretados) y “falsos negativos”(pronósticos de no delincuencia incumplidos)²⁵.

Diversos estudios llevados a cabo en Estados Unidos respecto de uno de los sistemas más difundidos en la Justicia Penal de ese país, denominado COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), han puesto en serio cuestionamiento diversos aspectos del mismo, tales como su eficacia, su presunta neutralidad, y su transparencia.

Con relación a la *eficacia*, se afirma que sus predicciones no se diferencian de valoraciones que pudieran efectuar ciudadanos comunes u otros algoritmos más sencillos en los mismos casos que el sistema COMPAS²⁶.

Respecto de su *neutralidad*, se han detectado notorios sesgos racistas²⁷, partiendo del hecho que la misma en su intento de predecir el riesgo de reincidencia utiliza como base un análisis complejo de información obtenida en base a 137 preguntas, cuyo contenido alude a aspectos que individuales que pueden favorecer la discriminación de ciertas personas o

²⁵ Ob. cit., pág. 102.

²⁶ Cita la autora el trabajo de Dressel y Farid, que informa que se realizó un estudio en el que personas legas en Derecho, a partir de 7 parámetros, valoraban la posible reincidencia de un grupo de presos y el porcentaje de acierto (62%) fue muy similar al obtenido por COMPAS (65%). Adicionalmente, en un segundo experimento, se comparó el resultado de COMPAS – que valora 137 parámetros- con el de un algoritmo mucho más simple – que incorporaba solo la edad y el número de condenas previas- y el porcentaje de acierto fue muy similar al de COMPAS, incluso ligeramente superior (67%) (DRESSEL, J; FARID, H. The accuracy, fairness and limits of predicting recidivism. *Science advances*, n. 4, 2018).

²⁷ Aquí menciona Neira Pena (ob. cit.: ídem), el estudio desarrollado por DIETRICH, William; MENDOZA, Christina; BRENNAN, Tim. COMPAS Risk Scales: Demonstrating accuracy equity and predictive parity. Northpointe Inc. Research Department, 2016, que habría demostrado que el sistema, a pesar de no incluir la raza directamente como parámetro, presenta un sesgo racial, al cotejar los falsos positivos y los falsos negativos. Concretamente, de los negros catalogados con riesgo alto, un 45% no delinquiró, mientras que, de los blancos catalogados con el mismo nivel de riesgo, solo un 23.5% fueron falsos positivos. Por el contrario, los blancos que obtuvieron un porcentaje de bajo riesgo y luego delinquieron fueron un 47.7%, frente a un 28% de los negros.

colectivos²⁸, asignando por ejemplo a las personas afroamericanas un puntaje alto (a pesar que posteriormente no reincidieron) y por el contrario a personas caucásicas puntaje bajo (aunque luego sí reincidieron)²⁹.

Finalmente, la acusación de *opacidad* de COMPAS se centra en la reivindicación de la protección del algoritmo por el derecho al secreto comercial de sus creadores, lo que impediría a los jueces y a los encausados o presos sometidos a valoración conocer el funcionamiento del algoritmo, así como los datos de que se nutre y la ponderación que el sistema hace de los mismos³⁰.

Por otro lado, aparecen un conjunto de objeciones o dilemas de carácter esencialmente jurídico, que apuntan a la posible afectación a derechos y garantías constitucionales a través del uso de la IA en el segmento de las medidas cautelares penales.

En primer lugar, toda una línea de derechos fundamentales, vinculadas a la *intimidad* de la persona o la protección de sus datos personales.

En una segunda dirección, los ligados a las formas constitucionales del *debido proceso*, como el derecho a la defensa en juicio o tutela judicial efectiva. Es que la carencia de la información sobre la cual se sustenta la decisión judicial, importa tanto la imposibilidad de confrontarla o contradecir en el debate adversarial como al momento de evaluar los fundamentos de la misma para que puedan ser objeto de revisión por un tribunal superior, es decir, ejercer el derecho al recurso.

La doctrina ha sido crítica de esta pretensión de opacidad, cuestionando que puedan admitirse las denominadas “*black box*”, en tanto resulta inadmisibles que la protección del derecho al secreto empresarial del creador intelectual de la herramienta impida el conocimiento del algoritmo, del peso ponderado que se otorga a cada factor variable, o del tipo de datos de que este se nutre por parte de los operadores jurídicos y de los justiciables.

Desde esta perspectiva se cuestiona seriamente la sentencia dictada en el caso *Wisconsin v. Loomis*, ya que la misma vino a legitimar la falta de

²⁸ Cfr. ROMERO CASABONA, Carlos María, “Inteligencia artificial, predictividad y justicia penal” en *Dupuy-Corbalán (dirs.) Cibercrimen III*, Montevideo-Buenos Aires, BdeF, 2020.

²⁹ RIQUERT, ob. cit.: pág. 111, con referencia a un trabajo de DANESI.

³⁰ Cfr. NEIRA PENA, ob. cit.: pág. 1916.

transparencia del algoritmo, con base en el derecho al secreto empresarial de la entidad que desarrolló COMPAS³¹.

Asimismo, también en el marco del proceso penal, y específicamente de la justicia cautelar, se encontrarían en contradicción derechos fundamentales de primer orden como la *presunción de inocencia* del encausado o su derivado principio de libertad personal durante el proceso, especialmente en aquellos casos en que esta última fuese limitada en atención a un perfil genérico de riesgo del sujeto en cuestión, sin atender suficientemente a sus circunstancias personales y a la necesidad de tratarle como inocente mientras no exista sentencia firme de condena³².

Por ello, en las reflexiones finales, recapitularemos sobre los principales límites que presentan estas herramientas, en el específico ámbito de las decisiones que atañen al encarcelamiento preventivo, como uno de los espacios más problemáticos del sistema penal.

4. Reflexiones finales: Límites de una Justicia Algorítmica en el campo de la prisión provisional.

Como cierre en las reflexiones finales, corresponde dar cuenta de los interrogantes formulados al inicio de la ponencia.

Por un lado, a) *si en las valoraciones judiciales de riesgo procesal al momento de adoptar medidas cautelares en el curso del proceso penal -en particular la posibilidad de la prisión provisional o preventiva- es admisible acudir al uso de programas o herramientas de IA.*

Por otro, b) *si en esa labor tales mecanismos resultan eficaces en términos de la esperada objetividad, certeza y reducción del margen de error que debería brindar la sofisticación de los recursos tecnológicos.*

³¹ NEIRA PENA, cita en tono crítico al respecto las posiciones de autoras como Monserrat De Hoyos Sancho y de Lucía Martínez Garay. Esta última cuestiona la decisión judicial aludida indicando que “para poder cuestionar con una mínima eficacia el resultado del algoritmo, el acusado detiene que poder saber cómo se ha llegado a él” por lo que, en su opinión, “los tribunales no deberían admitir valoraciones de riesgo hechas con algoritmos secretos”. (Ob. cit., pág. 1921, nota 60).

³² Cfr. NEIRA PENA, pág. 1919, nota 53.

En relación al primer dilema, y el modo en que -según se describió- pueden ser objeto de afectación derechos constitucionales básicos, corresponde ser cautos y prudentes en su incorporación a los sistemas de administración de justicia.

En cualquier caso, como sugiere la doctrina³³, debieran tenerse en consideración estricta los parámetros fundamentales de legalidad y previsibilidad de la decisión judicial a adoptarse con uso de estos mecanismos de IA. Asimismo, debe conjugarse adecuadamente su utilización con el respeto del derecho a la libertad respecto de quien, ha de ser considerado y tratado como inocente.

Por otra parte, debe garantizarse en forma amplia el esencial derecho de defensa del justiciable, asegurando la debida contradicción y las posibilidades de impugnar de forma efectiva la decisión cautelar.

Finalmente, no dejar de considerar las circunstancias concretas y específicas del caso y la persona involucrados, de modo de proporcionar al justiciable un trato justo, esto es, no discriminatorio o sesgado que, por lo tanto, responda a la necesaria neutralidad de una decisión imparcial.

Con relación al segundo de los interrogantes planteados, las investigaciones presentadas a la fecha siguen presentando resultados ambiguos, por lo que no parecen sugerir una necesidad imperiosa o ineludible de su implementación, máxime si se consideran los costos éticos de su utilización en términos de posibles afectaciones a derechos fundamentales.

Tampoco, desde un punto de vista práctico, las aspiraciones de lograr un impacto positivo en términos de reducción cuantitativa del encarcelamiento preventivo, sugieren un escenario irrefutable ni mucho menos, por el contrario numerosos estudios informan impactos de signo contrario, tales como los que habrían derivado en su momento, de los criterios de *justicia actuarial*.

En suma, la incorporación de la IA en este campo de la cautela procesal penal, debe ser cauta, paulatina y acotada, sin dejar encandilarse con la seducción de la innovación tecnológica.

³³ Se siguen en este punto las postulaciones formuladas por NEIRA PENA, ob. cit. pág. 1920.

En esa dirección, los sistemas de IA en ningún caso, pueden sustituir la necesaria decisión humana, tanto en relación a la administración de justicia penal en general como a las medidas de cautelar provisional durante el curso del proceso.

Es decir, su función debe reducirse, eventualmente, a un uso acotado exclusivamente como asistencia o apoyo a las decisiones judiciales, es estos los llamados sistemas *human-in-command* (HIC), cuya característica central resulta la de mantener la necesaria intervención humana en la decisión final³⁴.

En definitiva, esta postulación no resulta más que la *derivada del principio constitucional de estricta jurisdiccionalidad* que rige el debido proceso en general y las medidas cautelares, y especialmente las de mayor lesividad como la prisión preventiva.

Por tanto, la persistencia de esos valores constitucionales, que resultan fronteras infranqueables, exigen la satisfacción del requisito de *motivación* exhaustiva de las decisiones sobre medidas de encarcelamiento preventivo o atentatorias de otros derechos, de modo que en cualquier caso sea posible establecer un adecuado contralor del peso o la influencia que, en el marco de otros tantos elementos convictivos, pudiera tener el juicio algorítmico sobre las merituaciones efectuadas al momento de adoptar aquellas decisiones judiciales³⁵.

En suma, la aspiración a una Justicia Algorítmica, presenta a mi modo de ver obstáculos insalvables, centrados esencialmente en la propia naturaleza humano del complejo acto de juzgar, que contiene definiciones y valoraciones insusceptibles de ser replicadas por sistemas de IA; por lo que en cualquier caso estas aportaciones coadyvantes a la tarea de los actores del campo de la Justicia Penal serán únicamente admisibles en el marco de estrictos límites éticos y constitucionales.

³⁴ Cfr. NEIRA PENA, ob. cit. pág. 1922.

³⁵ Idem, pág. 1923, quien concluye que: “...En definitiva, para que las decisiones que afectan a la libertad del individuo puedan ser consideradas respetuosas con los derechos procesales fundamentales y, en última instancia, con la dignidad humana es absolutamente imprescindible que ... sean tomadas respetando el principio de legalidad y asegurando la transparencia del proceso valorativo y decisonal, el cual ha de desarrollarse de modo individualizado, atendiendo a las circunstancias particulares del caso concreto y de quienes sufren tales restricciones de derechos, lo que implica, en ultimo termino, que la decisión final, más o menos influida por la máquina, ha de ser tomada por aquel individuo en que se ha depositado la potestad jurisdiccional de forma exclusiva y excluyente, que no es otro que el juez...”. (NEIRA PENA, ob. cit. pág. 1927).

