


# Inteligencia artificial y tutela cautelar. Especial referencia a la prisión provisional

*Artificial intelligence and precautionary guardianship.  
Special reference to pre-trial detention*

**Ana María Neira Pena<sup>1</sup>**

Universidade da Coruña, España

ana.neira@udc.es

 <http://orcid.org/0000-0002-6562-8208>

---

**RESUMEN:** Esta investigación se centra en el estudio del uso de la IA en la toma de decisiones cautelares. Concretamente se examina la forma en que habitualmente los jueces valoran los presupuestos de la prisión provisional y se cuestiona si los sistemas de IA pueden replicar tales procesos valorativos, mejorando su eficacia. A tal fin, se analizan los sesgos en que incurren los jueces a la hora de decidir sobre la prisión provisional, evaluando sus diferentes presupuestos de adopción, y el modo en que las herramientas de IA podrían replicar y, en su caso, mejorar tales decisiones. Por último, se reflexiona sobre los riesgos que la introducción de sistemas de IA podría suponer para los derechos y garantías procesales, tratando de establecer las cautelas o salvaguardas precisas que debieran de flanquear el avance de la IA en la justicia penal para lograr una incorporación de la ciencia en la justicia que resulte respetuosa con los derechos fundamentales del justiciable. En definitiva, este trabajo pivota sobre dos cuestionamientos clave a los que se trata de dar respuesta: 1º) ¿El uso de la IA puede servir para mejorar la eficacia del proceso de toma de decisión en cuanto a la adopción de la prisión provisional?; 2º) ¿Qué riesgos plantea el uso de la IA en este contexto para los derechos y garantías del justiciable y qué cautelas deben de adoptarse para prevenir tales riesgos?

**PALABRAS CLAVE:** prisión provisional; inteligencia artificial; valoración de riesgos; garantías procesales.

---

<sup>1</sup> Profesora contratada doctora en el área de Derecho Procesal de la Universidade da Coruña (España). Doctora en Derecho por la Universidade da Coruña.

**ABSTRACT:** *This research focuses on the study of the use of AI in pretrial detention decision making. Specifically, it analyzes the way in which judges usually assess the requirements of pre-trial detention and questions whether AI systems can replicate such assessment processes, improving the efficiency of decision making. To this end, we analyze the biases that judges incur when deciding on pre-trial detention, assessing the purposes and dangers that justify it, and how AI tools could replicate and, when possible, improve such decisions. Finally, the paper reflects on the risks that the introduction of AI systems could pose to procedural rights and guarantees, trying to establish the precise precautions or safeguards that should flank the advance of AI in criminal justice to achieve an incorporation of science in justice that is respectful of the fundamental rights of the defendant. In short, this paper hinges on two key questions to be answered: 1<sup>st</sup>) Can the use of AI serve to improve the efficiency of the decision-making process regarding the adoption of pre-trial detention? 2<sup>nd</sup>) What risks does the use of AI in this context pose to the rights and guarantees of the defendant and what precautions should be taken to prevent such risks?*

**KEYWORDS:** *pre-trial detention; artificial intelligence; risk assessment; procedural safeguards.*

**SUMARIO:** *Introducción; 1. La decisión judicial cautelar: heurísticos y sesgos 2. El uso de la IA para valorar los presupuestos de la prisión provisional; 2.1. *Fumus commisi delicti*; 2.2. *Periculum in mora* o *periculum in libertatis*; 3.- Riesgos para los derechos fundamentales procesales y garantías a adoptar en el uso de la IA en la toma de decisiones cautelares; Conclusiones; Referencias.*

---

## INTRODUCCIÓN

La utilización de algoritmos e inteligencia artificial (en adelante, IA) en el ámbito judicial, y especialmente en la justicia penal<sup>2</sup>, es una

---

<sup>2</sup> Por destacar tan solo algunas publicaciones recientes en lengua castellana, véanse las siguientes: MIRÓ LLINARES, Fernando. *Inteligencia artificial y Justicia Penal: Más allá de los resultados lesivos causados por robots*. *Revista de Derecho Penal y Criminología*, n. 20, pp. 87-130, 2018; NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*. Madrid: Marcial Pons, 2018; PÉREZ

realidad en auge<sup>3</sup>. Las aplicaciones que la IA puede tener en el ámbito de la justicia penal son muchas y muy variadas, abarcando ámbitos tan diversos como el apoyo a la investigación policial, la prevención delictiva,

---

ESTRADA, Miren Josune. El uso de algoritmos en el proceso penal y el derecho a un proceso con todas las garantías. En BARONA VILAR, Silvia (ed.) *Claves de la justicia penal: feminización, inteligencia artificial, supranacionalidad y seguridad*. Valencia: Tirant lo Blanch, 2019, pp. 235-154; HERNÁNDEZ GIMÉNEZ, María. Inteligencia artificial y Derecho Penal. *Actualidad Jurídica Iberoamericana*, n. 10 bis, pp. 792-843, junio 2019; BORRÁS ANDRÉS, Núria. La verdad y la ficción de la inteligencia artificial en el proceso penal. En: CONDE FUENTES, Jesús; SERRANO HOYO, Gregorio (dirs.). *La justicia digital en España y la Unión Europea: Situación actual y perspectivas de futuro*. Barcelona: Atelier, 2019, pp. 31-39; GASCÓN INCHAUSTI, Fernando. Desafíos para el proceso penal en la era digital: externalización, sumisión pericial e inteligencia artificial. En CONDE FUENTES, Jesús; SERRANO HOYO, Gregorio (dirs.). *La justicia digital en España y la Unión Europea: Situación actual y perspectivas de futuro*. Barcelona: Atelier, 2019, pp. 191-206; BORGES BLÁZQUEZ, Raquel. El sesgo de la máquina en la toma de decisiones en el proceso penal. *Ius et scientia*, v. 6, n. 2, pp. 54-71, 2020; BALCELLS, Marc. Luces y sombras del uso de la inteligencia artificial en el sistema de justicia penal. En: CERRILLO I MARTÍNEZ, Agustí; PEQUERA POCH, Miquel (coords.). *Retos jurídicos de la inteligencia artificial*. Cizur Menor (Navarra): Aranzadi, 2020, pp. 145-159; SOLAR CAYÓN, José Ignacio. Inteligencia artificial en la justicia penal: los sistemas algorítmicos de evaluación de riesgos. En: SOLAR CAYÓN, José Ignacio (ed.). *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho*. Alcalá de Henares: Universidad de Alcalá de Henares, 2020, pp. 125-172; MUÑOZ RODRÍGUEZ, Ana Belén. El impacto de la inteligencia artificial en el proceso penal. *Anuario de la Facultad de Derecho. Universidad de Extremadura*, n. 36, pp. 695-728, 2020; ROMEO CASABONA, Carlos María. Inteligencia artificial, derechos fundamentales y proceso penal. *Comunicaciones en propiedad industrial y derecho de la competencia*, n. 89 (enero-abril), pp. 253-271, 2020.

<sup>3</sup> Prueba de ello es la preocupación que este fenómeno ha despertado, no solo en la doctrina, sino también en las instituciones. Como botón de muestra del interés que despiertan las aplicaciones de IA en la justicia puede verse la *Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno*, aprobada por la Comisión europea para la eficacia de la justicia el 3 de diciembre de 2018 (Estrasburgo). Disponible en <https://campusialab.com.ar/wp-content/uploads/2020/07/Carta-e%CC%81tica-europea-sobre-el-uso-de-la-IA-en-los-sistemas-judiciales-.pdf>. Último acceso: 03/09/2021) Esta Carta es la precursora del *Libro Blanco sobre la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza* (Bruselas, 19.2.2020 COM (2020) 65 final). Disponible en: [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_es.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf). Último acceso: 03/09/2021.

la valoración de las pruebas o la evaluación de riesgos<sup>4</sup>. Pero, de entre todas ellas, la denominada justicia predictiva<sup>5</sup>, especialmente cuando se usa para apoyar decisiones judiciales restrictivas de derechos – tales como la adopción de medidas de seguridad, penas o medidas cautelares – resulta especialmente problemática por los riesgos que acarrea para las los derechos fundamentales y las garantías procesales.

Esta investigación se centra en el estudio del uso de la IA en las decisiones cautelares, concretamente en la prisión provisional<sup>6</sup>, por ser la medida cautelar personal más restrictiva para los derechos del encausado y, por lo tanto, la que debe encontrarse más rodeada de garantías a la hora de adoptarse. En este sentido, el trabajo trata de determinar en qué medida el uso de la IA podría mejorar la eficacia en la toma de decisiones sobre la situación cautelar personal del encausado, a través de la valoración de las circunstancias que justifican la idoneidad y la necesidad de una medida de prisión provisional.

En la fase cautelar se trabaja en un contexto de relativa incertidumbre, en el que resulta esencial la valoración de riesgos, entorno en el que las potencialidades de la IA resultan indiscutibles<sup>7</sup>. Sin embargo,

---

<sup>4</sup> Sobre las diversas aplicaciones de la IA, y más en general de los avances tecnológicos, al proceso penal, véase HERNÁNDEZ GIMÉNEZ, María, op. cit., pp. 810-837.

<sup>5</sup> Sobre este uso de la IA, véase KEHL, Danielle; GUO, Priscilla; KESSLER, Samuel. *Algorithms in the Criminal Justice System: assessing the use of risk assessment in sentencing*. Responsive Communities Initiative, Berkman Klein Center for Internet & Society. Harvard Law School, 2017.

<sup>6</sup> Esta medida es definida como una medida cautelar penal, provisional y de duración limitada, que puede dictar el Juez mediante un auto especialmente motivado, por la que se restringe el derecho a la libertad del imputado de la comisión de un delito de especial gravedad y en quien concurra un peligro de fuga suficiente para presumir racionalmente que no acudirá al acto del juicio oral, destinada a asegurar dicha comparecencia, así como a conjurar los riesgos de reiteración delictiva, de ocultación o destrucción de las fuentes de prueba o la puesta en peligro de la integridad de la víctima (GIMENO SENDRA, Vicente. *Manual de derecho procesal penal*. 2ª ed. Madrid: Ediciones Jurídicas Castillo de Luna, 2018, p. 483)

<sup>7</sup> Sobre los sistemas de IA para la valoración de riesgos, véase SIMÓN CASTELLANO, Pere. *Justicia cautelar e inteligencia artificial*. La alternativa a los atávicos heurísticos judiciales, Barcelona: Bosch, 2021, pp. 130-148, quien explica de forma accesible cómo funcionan los sistemas de valoración de riesgo,

el uso de estas herramientas tecnológicas puede resultar especialmente controvertido si se toma en cuenta, en primer lugar, que hablamos de medidas restrictivas del derecho a la libertad y, en segundo lugar, que tales medidas se aplican a un sujeto que debe de ser considerado inocente, en tanto que aún no ha sido enjuiciado y condenado. Por eso el trabajo también trata de reflexionar sobre la forma en que la introducción de la IA como herramienta para la toma de decisiones cautelares privativas de libertad puede afectar a los derechos del justiciable, haciendo especial hincapié en las cautelas que deberían adoptarse para prevenir la vulneración o el menoscabo de tales derechos.

El objeto de este trabajo es, en definitiva, determinar en qué medida los sistemas de IA son aptos y eficaces para valorar la concurrencia de los presupuestos de la prisión provisional y de qué forma su uso puede ser introducido de manera respetuosa con los derechos del encausado. A tal fin, se analiza cómo son tomadas las decisiones cautelares por los jueces para, a continuación, determinar en qué medida sus procesos mentales son replicables, y en su caso mejorables, por sistemas de IA, capaces de gestionar ingentes cantidades de datos. A continuación, se exponen los principales riesgos que conllevaría el uso de tales sistemas en la justicia cautelar y las salvaguardas a adoptar para prevenirlos.

## 1. LA DECISIÓN JUDICIAL CAUTELAR: HEURÍSTICOS Y SESGOS

Para determinar en qué medida la IA puede ser útil en la toma de decisiones cautelares y, concretamente, para determinar la procedencia de la prisión provisional, primeramente, es necesario indagar sobre cómo se vienen desarrollando los procesos mentales de toma de decisiones en este ámbito, esto es, en sede judicial. En este sentido, la primera circunstancia a considerar, en tanto que determinante de la toma de decisiones cautelares, tiene que ver con el entorno de incertidumbre que

---

especialmente en el ámbito del *compliance* empresarial; Sobre los retos y los problemas que plantea la justicia predictiva desde el punto de vista del principio de legalidad, la presunción de inocencia o el derecho de defensa, véanse las reflexiones de ZAVRŠNIK, Aleš. Algorithmic justice: Algorithms and big data in criminal justice setting. *European Journal of criminology*, pp. 1-20, pp. 13-14, septiembre de 2019.

rodea la adopción de medidas cautelares, en una fase procesal en la que aún no se han producido ni practicado pruebas –que, en sentido estricto, y como regla general, sólo se producen en la fase de juicio oral- y en la que, sin embargo, se deben adoptar decisiones de enorme relevancia con base en indicios o elementos de convicción no siempre sólidos.

Tal y como aprecia Simón Castellano<sup>8</sup> la influencia del contexto más cercano al juzgador, junto con variables emocionales, condicionadas por aspectos como la apariencia o el lenguaje empleado por el encausado, tienen todavía más peso cuando se trata de decisiones que han de adoptarse de forma rápida y con una actividad probatoria nula o muy limitada, como ocurre con las medidas cautelares. Y, en este sentido, la ventaja que parece aportar la IA, en tanto que método automatizado de valoración de riesgos basado en evidencias y datos objetivos es, precisamente, el avance hacia la superación de los sesgos existentes en los juzgadores<sup>9</sup>. Sin embargo, no se puede menospreciar el hecho de que tanto las bases de datos como los algoritmos son creaciones humanas y, como tal, los modelos matemáticos en que se sustentan reflejan objetivos e ideologías<sup>10</sup>.

<sup>8</sup> SIMON CASTELLANO, Pere, op. cit., p. 120.

<sup>9</sup> Es lo que BORGES BLÁZQUEZ, Raquel, op. cit., p. 57, 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.

<sup>10</sup> O’NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown, 2016; En la misma línea, MIRÓ LLINARES, Fernando. “El modelo policial que viene: mitos y realidades del impacto de la inteligencia artificial y la ciencia de datos en la prevención policial del crimen”. En MARTÍNEZ ESPASA, José (coord.). *Libro blanco de la prevención y seguridad local valenciana: Conclusiones y propuestas del Congreso Valenciano de Seguridad Local: la prevención del siglo XXI, celebrado en Benidorm del 16 al 18 de noviembre de 2018*, Valencia: Agència Valenciana de Seguretat i Resposta a les Emergències: Agència Valenciana de Seguretat i Resposta a les Emergències, Instituto Valenciano de Seguridad Pública y Emergencias, 2019, pp. 98-113, pp. 102-103, sostiene que es un mito considerar que la IA se alimenta solo de números y complejas fórmulas matemáticas, ya que con ello se obvia el hecho de que “la IA no se apoya en números sino en datos, esto es, en información sobre la vida real, en información relacionada con variables concretas que interactúan entre sí a partir de modelos teóricos determinados que, en el caso de que tengan que ver con las personas, sólo pueden construirse sobre la base de Ciencias sociales como la Psicología, la Sociología o la Criminología. Es decir, aunque la IA requiere de análisis matemáticos, los mismos se realizan

Tanto Simón Castellano<sup>11</sup> como Nieva Fenoll<sup>12</sup> se refieren a los heurísticos de pensamiento descritos por Tversky y Kahneman<sup>13</sup> para describir algunos de los procesos mentales que siguen los jueces para tomar sus decisiones. Se distinguen, en este sentido, el heurístico de la representatividad<sup>14</sup>, el heurístico de accesibilidad o disponibilidad<sup>15</sup>, el heurístico de anclaje o ajuste<sup>16</sup> y el de afección<sup>17</sup>.

De acuerdo con el heurístico de representatividad, el enjuiciamiento humano juzga la probabilidad de una hipótesis acudiendo a los datos de su memoria, por lo tanto, acudiendo a su experiencia previa en situaciones análogas y al éxito de sus decisiones anteriores. A través de este heurístico, un suceso se considera probable en la medida en que representa las

---

*sobre variables sociales, y son los científicos sociales quienes deben configurar los marcos explicativos adecuados para ello”.*

<sup>11</sup> SIMÓN CASTELLANO, Pere, op. cit., pp. 116-120.

<sup>12</sup> NIEVA FENOLL, Jordi, op. cit., pp. 45-60.

<sup>13</sup> KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982.

<sup>14</sup> KAHNEMAN, Daniel; TVERSKY, Amos, “Judgments of and by representativeness”. En KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982, pp. 84-98, esp. p. 97. De acuerdo con los referidos autores, los resultados de los estudios realizados apoyan directamente la hipótesis de que las personas evalúan la probabilidad de los sucesos en función del grado en que estos son representativos de un modelo o proceso relevante. Dado que la representatividad de un acontecimiento puede incrementarse mediante la especificidad, un objetivo compuesto puede juzgarse más probable que uno de sus componentes.

<sup>15</sup> De acuerdo con KAHNEMAN, Daniel; TVERSKY, Amos. Availability: A heuristic for judging frequency and probability. En KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982, pp. 163-178, se dice que una persona emplea la heurística de la disponibilidad siempre que estima la frecuencia o la probabilidad por la facilidad con la que se pueden traer a la mente ciertos eventos o asociaciones.

<sup>16</sup> El heurístico de anclaje o ajuste supone, en el ámbito de las decisiones judiciales, que el juzgador tiende a hacerse una idea inicial del asunto y, posteriormente, a reinterpretar los datos en concordancia con tal posicionamiento inicial, muchas veces tamizado por su personal forma de entender e interpretar la realidad.

<sup>17</sup> La afección se refiere a la influencia que, en las decisiones judiciales, tienen las emociones.

características esenciales de su población de origen o proceso generador de acuerdo con la experiencia propia del decisor. Por su parte, el heurístico de accesibilidad se basa en el proceso mental por el cual el ser humano otorga una mayor probabilidad a aquellas hipótesis que recuerda con mayor facilidad, esto es, aquellas más accesibles en su memoria. El heurístico de anclaje o ajuste supone, en el ámbito de las decisiones judiciales, que el juzgador tiende a hacerse una idea inicial del asunto y, posteriormente, a reinterpretar los datos en concordancia con tal posicionamiento inicial, muchas veces tamizado por su personal forma de entender e interpretar la realidad. Por último, la afección influye igualmente en la toma de decisiones introduciendo el componente emocional, aspecto que, ya se puede adelantar, parece, al menos en la actualidad, de imposible replicación por los sistemas de IA. Aunque, respecto de este último, la pregunta de más difícil respuesta pasaría seguramente por determinar en qué resulta beneficioso y recomendable eliminar del proceso de decisión la empatía o la inteligencia emocional<sup>18</sup>.

De la conjunción de los referidos procesos psicológicos de toma de decisiones, se deriva el hecho de 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 podría 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. Así pues, tal y como advierte Simón Castellano<sup>19</sup>, 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*

---

<sup>18</sup> Tal y como indica ZAVRŠNIK, Aleš, op. cit., p. 12, la empatía u otras cualidades personales son tipos de sesgos que mejoran la equidad medible estadísticamente, por lo que quizás no deberían de eliminarse del ámbito judicial.

<sup>19</sup> SIMÓN CASTELLANO, Pere, op. cit., p. 122.



*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”.*

Se debe de advertir que los referidos procesos mentales, empleados por el juez cotidianamente en su toma de decisiones, no están exentos de sesgos<sup>20</sup>. La cuestión es determinar si tales formas de tomar decisiones pueden ser replicadas, y en su caso mejoradas, con sistema de IA que, manejando datos en mayor cantidad y de mejor calidad que los retenidos por la memoria humana, puedan servir de complemento eficaz a las decisiones judiciales, al nutrir la valoración judicial sobre los presupuestos de adopción de las medidas cautelares.

Ahora bien, siendo 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 algoritmo<sup>21</sup>.

---

<sup>20</sup> La existencia de estos sesgos ya fue advertida por KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos, op. cit., p. 20, en cuya opinión estos heurísticos son muy económicos y suelen ser eficaces, pero conducen a errores sistemáticos y predecibles. Una mejor comprensión de estos heurísticos y de los sesgos a los que conducen podría mejorar los juicios y las decisiones en situaciones de incertidumbre. Así, por ejemplo, el heurístico de accesibilidad podría dar lugar al sesgo de correlación ilusoria, acuñado por CHAPMAN, Loren James. Illusory correlation in observational report. *Journal of Verbal Learning & Verbal Behavior*, v. 6, n. 1, pp. 151–155, 1967. Este sesgo implica sobreestimar la existencia de covariación entre dos eventos y constituye un sesgo de confirmación, a través del cual el decisor tiende a basarse únicamente en casos confirmatorios de sus ideas o hipótesis, a la vez que ignora los casos no confirmatorios.

<sup>21</sup> En este sentido, tal y como indica MARTÍNEZ GARAY, Lucía. Peligrosidad, Algoritmos y *Due Process*: el Caso *State V Loomis*. *Revista de Derecho Penal y Criminología*, 3.<sup>a</sup> Época, n 20 pp. 485-502, p. 496, julio de 2018, “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”.

## 2.- EL USO DE LA IA PARA VALORAR LOS PRESUPUESTOS DE LA PRISIÓN PROVISIONAL

### 2.1.- *FUMUS COMMISI DELICTI*

La elevada probabilidad de que se haya cometido un delito y los indicios racionales que permitan atribuir tal delito al investigado constituyen el primer presupuesto de adopción de las medidas cautelares personales en general; y, tratándose de un delito grave, podrían constituir el presupuesto inicial de adopción de la medida de prisión provisional. Esta condición necesaria, que no suficiente, para acordar la prisión provisional se basa en la existencia de motivos bastantes para creer responsable criminalmente al encausado (art. 503.1.2º LECrim), en el sentido de que no basta con la mera concurrencia de indicios racionales de criminalidad, sino que ha de acreditarse, además, que no concurren causas de exención o de extinción de la responsabilidad penal<sup>22</sup>. Todo ello supone la necesidad de desarrollar una valoración conjunta de los elementos de convicción que obren en la causa, desde una perspectiva fáctica y jurídica, considerando a tal efecto las alegaciones que las partes realicen al respecto.

Por lo tanto, como presupuesto básico de adopción de las medidas cautelares, especialmente de la prisión provisional, habrá que valorar la existencia de indicios fundados de responsabilidad criminal en el sujeto en cuestión, sin que ese juicio anticipado sobre su participación en los hechos y sobre su responsabilidad personal pueda confundirse o contaminarse con la estimación del nivel de predisposición a delinquir o con su actitud obstruccionista frente a la investigación, entre otros extremos no directamente vinculado con la comisión del delito que está siendo investigado.

Pues bien, hay que advertir que, en este ámbito, la ayuda que los sistemas de IA podrían brindar es limitada, más allá de sus potencialidades como instrumento de apoyo a la investigación policial<sup>23</sup> o de apoyo a la

---

<sup>22</sup> GIMENO SENDRA, Vicente, op. cit., p. 487.

<sup>23</sup> A este respecto pueden resultar útiles instrumentos como VALCRI (“*Visual Analytics for sense making in Criminal Intelligence Analysis*”, que sirve a los investigadores para relacionar evidencias obtenidas en la escena del crimen con datos contenidos en bases de datos policiales, con ayuda de técnicas

valoración de pruebas u otros elementos de convicción<sup>24</sup>. En tanto que la valoración de este presupuesto implica reconstruir hechos pasados, atribuirles un nivel de verosimilitud y valorarlos jurídicamente, las herramientas de justicia predictiva o valoración de riesgos no deben de influir en absoluto en este juicio, que debe de realizarse de forma independiente y autónoma a toda circunstancia ajena a la comisión del hecho delictivo que se investiga.

Por otra parte, aunque no se puede exigir certeza sobre la responsabilidad penal del investigado debido a la fase procesal en la que se adoptan normalmente las medidas cautelares, para acordar una medida de prisión provisional, el juez ha de estar convencido de que el delito se ha cometido, con un grado de probabilidad muy elevado<sup>25</sup>. Además, aunque la distribución de probabilidad de los diversos modelos permita cuantificar el grado de incertidumbre, esto nunca eliminará la necesidad de fijar un umbral, esto es, de determinar cuánto peso debe de darse a la certeza del riesgo esperado, trasladando la probabilidad algorítmica al lenguaje judicial<sup>26</sup>. Como consecuencia, será el juez, en último término, el

---

biométricas y de reconocimiento facial. Sobre este sistema, véase ESTÉVEZ MENDOZA, Lucana. Algoritmos policiales basados en IA derechos fundamentales a la luz de Hart y Valcri. En JIMÉNEZ CONDE, Fernando; BELLIDO PENADÉS, Rafael (coords.). *Justicia: garantías vs. Eficacia*. Valencia: Tirant Lo Blanch, 2019, pp. 665-674.

<sup>24</sup> En este ámbito resultan esenciales los avances y aportaciones de la psicología forense para valorar la credibilidad de los testigos, en combinación con herramientas automatizadas de tratamiento de datos que se nutren de la ciencia psicológica. Por poner un ejemplo de estas herramientas, cabe señalar VERIPOL, aplicativo introducido por la policía española en el año 2017 para detectar denuncias falsas de robos de móviles. El programa procesa el texto de la denuncia, extrayendo características útiles para su clasificación utilizando técnicas de procesamiento del lenguaje natural. Estas características se pasan a un modelo matemático que estima la probabilidad de falsedad de la denuncia. Experimentos empíricos muestran que el programa tiene una precisión superior al 90% ([http://www.interior.gob.es/prensa/noticias/-/asset\\_publisher/GHU8Ap6ztgsg/content/id/9496864](http://www.interior.gob.es/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864), último acceso: 03/09/2021)

<sup>25</sup> NIEVA FENOLL, Jordi, op. cit., p. 72, habla de la necesidad de determinar prácticamente una certeza de culpabilidad a través del enjuiciamiento de la culpabilidad de forma provisional. Hay que tener en cuenta, sin embargo, que, evidentemente, tal y como advierte ZAVRŠNIK, Aleš, op. cit., p. 10, los instrumentos de evaluación de riesgos arrojan probabilidades y no certezas.

<sup>26</sup> ZAVRŠNIK, Aleš, op. cit., p. 10.

que, valorando conjuntamente los elementos de convicción disponibles, entre los que podría estar el resultado algorítmico directo o indirecto acerca de la probabilidad de comisión del hecho, determine en qué medida considera suficientemente acreditado que el investigado ha cometido un delito de cierta gravedad.

El *fumus boni iuris* es, en definitiva, un juicio de imputación reforzado, esto es, un juicio provisional sobre la atribución de responsabilidad al encausado por un determinado hecho<sup>27</sup>. Se trata, por lo tanto, de reconstruir el pasado, valorando jurídicamente los hechos investigados, y no de predecir el futuro ni de asignar probabilidades a hechos inciertos. Y, en este ámbito, la IA es más limitada. Desde luego, difícilmente puede sustituir al juicio jurisdiccional, que además ha de guiarse por principios como el *in dubio pro reo* o el favorecimiento de la libertad personal, que encierran auténticos valores, no siempre reducibles a una secuencia lógica o a un umbral estadístico.

En definitiva, es importante hacer hincapié en que este extremo –la concurrencia del *fumus comissi delicti*– siga siendo valorado por el juzgador, y que no se confunda o se distorsione con valoraciones sobre el riesgo de reincidencia delictiva o con predicciones de comportamiento futuro, que lleven a confundir el juicio de imputación con el juicio de peligrosidad<sup>28</sup>.

La posibilidad de confusión se torna más probable si se toma en consideración que en la etapa cautelar existe, por definición, un vacío probatorio, que obliga al juez a decidir con una información incompleta, lo que podría incentivar al juzgador a guiarse por los perfiles de riesgo

---

<sup>27</sup> Este juicio de imputación se integra por un elemento objetivo, el hecho punible, y el subjetivo, su razonable atribución al destinatario de la medida en cuestión. Así pues, el *fumus comissi delicti* implica atribuir, fundadamente, la comisión de un hecho presuntamente delictivo a una determinada persona y exige valorar, por lo tanto, los elementos que conforman el objeto del juicio penal, esto es, “*si una persona determinada ha cometido o no una acción penal determinada de que deba responder criminalmente*” (GÓMEZ ORBANEJA, Eduardo. *Comentarios a la Ley de Enjuiciamiento Criminal*. Barcelona: Bosch, 1947, Tomo I., p. 252)

<sup>28</sup> Sobre el riesgo de confundir peligrosidad y autoría como consecuencia del uso de las herramientas de IA en las medidas cautelares advierte NIEVA FE-NOLL, Jordi, op. cit., p. 74.

producidos por el sistema de IA para fundamentar, incluso de forma inconsciente, el juicio de imputación<sup>29</sup>. Pues bien, en aras de la libertad y dignidad del investigado este riesgo debe de ser prevenido. Como cautela sería necesario que el órgano decisor no hiciese uso de los instrumentos de IA para predecir el riesgo de fuga o reiteración delictiva hasta haber apreciado suficientemente, de forma previa y autónoma, el *fumus commisi delicti*.

## 2.1.- *PERICULUM IN MORA O PERICULUM IN LIBERTATIS*

La IA se muestra, sin embargo, más útil para valorar el otro presupuesto procesal de las medidas cautelares: el *periculum in mora* o los riesgos concretos que las medidas cautelares tratan de precaver<sup>30</sup>. Aunque su uso en este ámbito en absoluto está exento de riesgos ni debe de estar desprovisto de cautelas, la potencialidad de la IA para fundamentar estas valoraciones resulta indudable<sup>31</sup>.

---

<sup>29</sup> Se debe de advertir, en este sentido, que el uso de sistemas de IA como complemento o apoyo a la toma de decisiones judiciales modifica la percepción de responsabilidad sobre las decisiones propias y, además, crea una inercia en el juzgador, que tiende a confiar en los resultados que ofrece el sistema, optando por justificarlos, acomodando sus propios razonamientos al resultado algorítmico (ZAVRŠNIK, Aleš, op. cit., p. 13). La misma idea se recoge en: GORDON, Faith. Book Review on EUBANKS, Virginia. Automating inequality: How high-tech tools profile, police, and punish the poor, *Law, Technology and Humans*, St Martin's Press, New York, v. 1, pp. 162-164, 2018.; GASCÓN INCHAUSTI, Fernando, op. cit., pp. 204-205; DE HOYOS SANCHO, Montserrat. El Libro Blanco sobre inteligencia artificial de la Comisión Europea: reflexiones desde las garantías esenciales del proceso penal como 'sector del riesgo'. *Revista Española de Derecho Europeo*, n. 76, pp. 9-44, p. 36, 2020.

<sup>30</sup> Véanse las reflexiones de NIEVA FENOLL, Jordi, op. cit., pp. 63-77 y de SIMÓN CASTELLANO, Pere, op. cit., pp. 149-157, sobre la forma en que la IA podría ser usada para valorar los diferentes riesgos que justifican la adopción de medidas cautelares.

<sup>31</sup> Las herramientas de evaluación de riesgos se encuentran, a día de hoy, en la cuarta generación de desarrollo, habiendo alcanzado un alto grado de sofisticación, empleando algoritmos de aprendizaje automático (*machine learning*) que generan modelos de riesgos basados en ingentes cantidades de datos (KEHL, Danielle; GUO, Priscilla; KESSLER, Samuel, op. cit. p. 9).

Los riesgos clásicos que se tratan de neutralizar con la prisión provisional son el riesgo de fuga, quizás el único propiamente cautelar<sup>32</sup>; el riesgo de reiteración delictiva, con una finalidad más preventiva que cautelar; y el riesgo de destrucción de pruebas.

Respecto del riesgo de fuga, se constata que la clásica finalidad cautelar de la prisión provisional consiste en *asegurar la presencia del investigado o encausado en el proceso cuando pueda inferirse racionalmente un riesgo de fuga* (art. 503.3º a) LECrim). Para valorar el referido riesgo, la propia legislación aporta una serie de indicadores, que deberán valorarse conjuntamente. Estos son: la naturaleza del hecho, la gravedad de la pena que pudiera imponerse al investigado o encausado<sup>33</sup>, la situación familiar, laboral y económica de este, así como la inminencia de la celebración del juicio oral<sup>34</sup>.

Sin embargo, la valoración conjunta de tales elementos conlleva sesgos. Así, por ejemplo, la sistemática discriminación de los extranjeros

---

<sup>32</sup> En esta línea, RAMOS MÉNDEZ, Francisco. *Enjuiciamiento criminal. Décima lectura constitucional*. Barcelona: Atelier, 2011, p. 255, sostiene que la necesidad de garantizar la sujeción de una persona al juicio penal es lo único que autoriza la medida cautelar de tipo personal.

<sup>33</sup> La jurisprudencia insiste en que *la gravedad de la pena* es un elemento clave para valorar el riesgo de fuga, dado que a mayor gravedad de la amenaza penal mayores serán los incentivos a la huida. Sin embargo, se afirma también que la gravedad de la pena no puede servir como único factor a tener en cuenta para acordar la prisión provisional, sino que deberá valorarse a la luz de las *circunstancias personales del encausado*, tales como el arraigo familiar, profesional y social, las conexiones con otros países y los medios económicos de los que este dispone, entre otros (STC 128/1995, de 26 de julio, FJ 4 b)).

<sup>34</sup> Respecto a la *proximidad de la celebración del juicio oral*, se afirma que, como norma general, cuando se acerca la fecha de celebración del juicio oral, la amenaza de la pena y la cercanía temporal de su posible imposición también operan como potenciales estímulos a la huida. Sin embargo, este dato presenta un sentido ambivalente o no concluyente, dado que el avance del proceso puede contribuir tanto a cimentar con mayor solidez la imputación, como a debilitar los indicios de culpabilidad del acusado, por lo que el órgano judicial debe concretar las circunstancias que avalan en el caso concreto una u otra hipótesis (SSTC 128/1995, de 26 de julio, FJ 4 b) y 66/1997, de 7 de abril, FJ 6, entre otras muchas). Esa naturaleza ambivalente también puede apreciarse en la circunstancia consistente en disponer de medios económicos, que puede ser vista como indicativa de posibilidades de urdir una fuga eficiente o, contrariamente, de falta de interés de separarse de su patrimonio (NIEVA FENOLL, Jordi, op. cit., p. 76).

en la valoración del riesgo de fuga como determinante de la adopción de la medida de prisión provisional ha sido denunciada tanto por la literatura española como por la extranjera a lo largo de los años<sup>35</sup>. Detrás de esta discriminación parece estar el sesgo de confirmación denominado correlación ilusoria<sup>36</sup>, propio del heurístico de representatividad, que conduciría a sobreestimar la correlación entre la condición de extranjero y las probabilidades de fugarse o de sustraerse a la acción de la justicia, subestimando, al mismo tiempo, otras circunstancias del caso que podrían abogar por la imposición de medidas cautelares menos restrictivas que la prisión.

---

<sup>35</sup> Véase, en este sentido, GUERRA PÉREZ, Cristina. *La decisión judicial de prisión preventiva*. Valencia: Tirant Lo Blanch, 2010, pp. 156-157; RECIO JUÁREZ, Matías. Nuevos Instrumentos para el Cumplimiento Transnacional de las Medidas Cautelares Alternativas a la Prisión Provisional en la Unión Europea. *Revista Direito e Inovação*, v. 2, n. 2, 198-213, pp. 200-204, 2014; GARCÍA ESPAÑA, Elisa. *Extranjeros sospechosos, condenados y ex condenados: Un mosaico de exclusión*. *Revista Electrónica de Ciencia Penal y Criminología*, n. 19-15, p. 19, 2017; LJUNGQUIST, Thomas. *Mutual Recognition of Non-Custodial Pre-Trial Supervision Measures in the European Union*. *Revue internationale de droit pénal*, v. 77, n. 1-2, pp. 172-173, 2006; BANACH-GUTIÉRREZ, Joanna Beata. *Globalized Criminal Justice in the European Union Context. How Theory Meets Practice*. *New Journal of European Criminal Law*, v. 4, n. 1-2, p. 165, 2013; RAFARACI, Tomasso. *The Application of the Principle of Mutual Recognition to Decisions on Supervision Measures as an Alternative to Provisional Detention*. En RUGGERI, Stefano (ed.). *Liberty and Security in Europe*. Osnabruck: Universitätsverlag, 2013, p. 68. Esta discriminación sistemática también se aprecia en el contexto de la Unión Europea. En este sentido, se afirma que los ciudadanos de la UE que no residen en el territorio del Estado miembro en el que son sospechosos de haber cometido un delito se encuentran a veces -principalmente debido a la falta de vínculos con la comunidad y al riesgo de fuga- en prisión preventiva. Una persona sospechosa de haber cometido un delito en un país en el que reside se beneficiaría, en una situación similar, de una medida de vigilancia menos coercitiva, como la obligación de presentarse periódicamente a la policía o la prohibición de viajar (Exposición de motivos de la Propuesta original de Decisión marco del Consejo sobre la orden europea de vigilancia en el marco de las medidas cautelares aplicadas entre los Estados miembros de la Unión Europea (COM (2006) 468 final), Bruselas, 29.8.2006).

<sup>36</sup> Sobre esta idea, véase JENNINGS, Dennis L.; AMABILE, Teresa M.; ROSS, Lee. Informal covariation assessment: Data-based versus theory-based judgment. En: KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982, pp. 211-230.

De acuerdo con Simón Castellano<sup>37</sup> la utilización de los sistemas de IA en este ámbito podría servir, no solo para replicar el proceso cognitivo humano en la valoración de tal riesgo, sino también para buscar patrones similares a través del volcado masivo de datos, haciendo aflorar resultados estadísticos sobre la probabilidad de que el investigado, en una causa concreta, acabe sustrayéndose a la acción de la justicia<sup>38</sup>. Sin embargo, esta forma de proceder plantea el problema de que podría conducir a automatismos, perpetuando los sesgos propios del razonamiento judicial en esta materia, o incluso agravándolos, al reproducirse patrones de decisión con base en decisiones ya adoptadas, sin reconsiderar el nivel de riesgo, y la proporcionalidad de la medida, de una forma verdaderamente individualizada y adaptada a las particularidades del caso concreto<sup>39</sup>. Es el heurístico de anclaje elevado a su máximo potencia: los datos de mayores tasas de prisión para extranjeros pueden realimentar y consolidar esta tendencia discriminatoria<sup>40</sup>.

---

<sup>37</sup> SIMÓN CASTELLANO, Pere, op. cit., , p. 78.

<sup>38</sup> De hecho, en un estudio de 1.36 millones de presos preventivos se observó que un sistema de IA podía predecir si un sospechoso se fugaría o reiteraría la conducta delictiva mejor que un juez humano (KLEINBERG, Jon; LAKKA-RAJU, Himabindu; LESKOVEC, Jure; LUDWIG, Jens; MULLAINATHAN, Sendhil. Human Decisions and Machine Predictions. *The Quarterly Journal of Economics*, v. 133, n. 1, pp. 237–293, 2018)

<sup>39</sup> De hecho, en el conocido caso Loomis uno de los argumentos del recurrente fue, precisamente, que se había vulnerado su derecho a obtener una sentencia individualizada. Sin embargo, la Corte Suprema de Wisconsin (Sentencia 13 de julio de 2016: *State v. Loomis*, 881, N.W.2d 749, 7532 (Wis, 2016)) rechazó el argumento, señalando que el recurrente tendría razón si el sistema COMPAS hubiese sido el único elemento a considerar en la valoración del riesgo, dado que la herramienta valora el riesgo de reincidencia para *grupos* de personas y no para *individuos* concretos. Sin embargo, en tanto que la herramienta sea utilizada como un elemento más de la decisión judicial, y valorada junto con el resto de circunstancias particulares propias del caso y del concreto individuo, su uso es legítimo. Sobre esta resolución, véanse las acertadas críticas de MARTÍNEZ GARAY, Lucía, op. cit., pp. 490-502.

<sup>40</sup> Algo similar es apreciado por DE HOYOS SANCHO, Montserrat, op. cit., p. 28, en relación con el uso que el sistema COMPAS hace del dato sobre el número de arrestos previos de personas con una determinada característica personal como es la raza o etnia. Como consecuencia de la recolección de este dato, por el mero hecho de ser afroamericano, el nivel de riesgo que indicará el algoritmo se multiplicará notablemente, ya que en Estados Unidos la cifra de detenciones de personas afroamericanas, y de otras minorías, es muy



Para que el uso de sistemas de IA suponga una verdadera ventaja a la hora de valorar el riesgo de fuga es importante, por una parte, tratar de compensar los referidos sesgos – y otros de género, edad, tipología delictiva, etc.- en la construcción del algoritmo, en la selección de los datos y, especialmente, en la interpretación contextual de los resultados que el sistema arroje; y, por otra parte, que el sistema se nutra y maneje datos suficientes en cantidad y, sobre todo, de buena calidad, tanto de casos anteriores, como especialmente del caso concreto<sup>41</sup>. Por lo tanto, un formulario que recoja debidamente toda la información relevante del caso y la sistematice puede resultar de gran relevancia para un proceso eficaz de toma de decisiones.

Otra de las finalidades de la prisión provisional prevista en la ley consiste en evitar la reiteración delictiva o ataques contra bienes jurídicos de la víctima. La legitimidad de esta finalidad, a pesar de presentar un carácter más preventivo que propiamente cautelar, ha sido avalada por el TEDH<sup>42</sup> y por el TC español<sup>43</sup>. Sin embargo, es necesario advertir que, conforme a la jurisprudencia del TEDH, dicha previsión no da cobertura

---

elevada. Sin embargo, tal y como aprecia la autora, el hecho de que se produzca ese mayor número de contactos entre la policía y los sujetos pertenecientes a minorías étnicas, no significa necesariamente una mayor propensión delictiva de estos, sino más bien que existe un control policial mucho más estricto sobre los sujetos pertenecientes a esos grupos raciales minoritarios en Estados Unidos.

<sup>41</sup> Así, por ejemplo, para evaluar el arraigo económico de una persona será relevante valorar si tiene trabajo, pero también el tipo de trabajo, el tipo de contrato, los años de antigüedad, el salario percibido, en su caso la posibilidad de reincorporarse al trabajo tras cumplir la pena, las relaciones con sus compañeros, etc. En este sentido, SIMÓN CASTELLANO, Pere, op. cit., p. 151, propone la incorporación al sistema de todos los datos que el encausado aporte voluntariamente (datos de navegación, correo electrónico, etc.). Esta propuesta resulta, a mi juicio, un tanto controvertida, ya que, aunque la aportación sea voluntaria, plantea el problema de la diferente fiabilidad del resultado del algoritmo según el grado de colaboración del investigado, así como el riesgo de que los datos presentados por el investigado sean sesgados o parciales.

<sup>42</sup> SSTEDH de 6 de noviembre de 1980, caso Guzzardi c. Italia; de 7 de marzo de 2013, caso Ostendorf c. Alemania, y 28 de octubre de 2014, caso Urtans c. Letonia.

<sup>43</sup> SSTC 128/1995, de 26 de julio, FJ 3º y 27/2008, de 11 de febrero, FJ 4º, entre otras.

a decisiones de prevención general dirigidas contra un individuo, o una categoría de individuos, considerados peligrosos debido a su continua tendencia al crimen; sino que, más limitadamente, en el contexto de la persecución de un delito, los arts. 5.1 c) y 5.3 CEDH, interpretados conjuntamente, permiten a los Estados contratantes imponer y mantener en el tiempo una privación cautelar de libertad previa al juicio como medio de prevención de una concreta y específica infracción penal, finalidad que ha de venir fundamentada en hechos o informaciones concretas basadas en datos objetivos<sup>44</sup>. Y, precisamente con el fin de objetivar al máximo el juicio de inferencia sobre el riesgo de reiteración delictiva, se viene exigiendo que entre el hecho presuntamente delictivo y el que se pretende evitar exista una naturaleza común, que ambos sean de cierta gravedad y que se valoren las posibilidades reales del encausado – atendiendo a su experiencia, capacidad y disposición- de reiterar la conducta<sup>45</sup>.

El problema que presenta el uso de la IA para la valoración de este extremo tiene que ver con el hecho de que el mismo concepto de “reiteración delictiva” parte de asumir como cierto que se ha cometido un delito, cuando en realidad, en el contexto de la adopción de la prisión provisional, o de otras medidas cautelares penales, el encausado debe de ser considerado inocente<sup>46</sup>. La valoración de este riesgo lleva implícita, por lo tanto, la apreciación positiva de la concurrencia del *fumus commissi delicti* que, como se indicó antes, no es un ámbito propicio para el uso de los análisis predictivos, en tanto que se refiere a la reconstrucción de hechos pasados<sup>47</sup>.

<sup>44</sup> SSTEDH de 6 de noviembre de 1980, caso Guzzardi c. Italia, § 102; de 7 de marzo de 2013, caso Ostendorf c. Alemania, §§ 67 a 69, y 28 de octubre de 2014, caso Urtans c. Letonia, § 33.

<sup>45</sup> SIMÓN CASTELLANO, Pere, op. cit., pp. 153-154.

<sup>46</sup> En este sentido, tal y como advierte NIEVA FENOLL, Jordi, op. cit., p. 67, “una cosa es establecer la posibilidad de reiteración delictiva de un reo condenado, es decir, con respecto al que ya se comprobó su responsabilidad, y otra muy distinta hacer idéntico juicio con alguien que aún no ha sido juzgado, es decir, con un inocente”. Por su parte, SIMÓN CASTELLANO, Pere, op. cit., pp. 154-155 indica que el uso de sistemas que predigan la reiteración delictiva en el ámbito de la justicia cautelar parte de un sesgo conceptual en el lenguaje y prejuiza antes del juicio.

<sup>47</sup> En este sentido, tal y como sostiene DE HOYOS SANCHO, Montserrat, op. cit., p. 32, nota 60, resulta complicado admitir juicios predictivos de peligrosidad

Es por ello que no faltan autores que sostengan que el uso de estas herramientas de predicción del riesgo de reiteración delictiva debería de limitarse a las decisiones propias de la fase de ejecución de condena, sin que deban ser usadas en materia cautelar, donde es posible que terminen por perpetuar, o incluso que tiendan a reforzar, los errores y sesgos preexistentes del juicio humano, con el hándicap añadido de que el juicio de peligrosidad –*periculum in mora*– puede llegar a subsumir al de autoría o culpabilidad –*fumus commissi delicti*–<sup>48</sup>. Otros, sin embargo, afirman que, si se admiten las “máximas de la experiencia” usadas en el razonamiento judicial para efectuar juicios predictivos en relación con las medidas cautelares, las cuales no dejan de estar basadas en la generalización de experiencias con otros sujetos, no tiene sentido limitar el uso de la IA en este ámbito, siempre que se realice con las debidas garantías<sup>49</sup>.

En todo caso, hay que tener presente que los sistemas de IA predictivos no están exentos de sesgos y de riesgos. Primeramente, porque el algoritmo o los datos con que aquel ha sido entrenado pueden tender a introducir o perpetuar las discriminaciones presentes en la actuación policial y judicial; y, en segundo lugar, porque su fiabilidad, en ocasiones, ha demostrado no ser mucho mayor que la de la intuición humana, aunque, en este caso, revestida de un científicismo, que dota a las decisiones de una fuerza categórica cuando menos preocupante<sup>50</sup>.

---

de un sujeto cuando todavía tiene la presunción de inocencia a su favor; aunque, a renglón seguido, la autora reconoce que, dado que es generalmente admitido el uso de las medidas cautelares con la finalidad de prevenir la reiteración delictiva, restringir el uso de la IA para valorar tal riesgo no tiene mucho sentido.

<sup>48</sup> NIEVA FENOLL, Jordi, op. cit., pp. 73-74.

<sup>49</sup> DE HOYOS SANCHO, Montserrat, op. cit., p. 32; Según GIALUZ, Mitja. Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa. *Diritto Penale Contemporaneo*, 2019, pp. 19-20., las limitaciones que se tratan de imponer al uso de las herramientas de justicia predictiva en materia cautelar no están justificadas, en tanto que no tiene sentido limitar los instrumentos cognitivos del juez. Además, el autor sostiene que la verdadera razón de tales limitaciones en la legislación italiana se explica por la poca confianza del legislador en la ciencia psicológica y criminológica.

<sup>50</sup> La herramienta estrella en el ámbito de la valoración de riesgo de reincidencia, aunque existen muchas otras, es sin duda, COMPAS, que se utiliza en

Como cautelas ante el previsible uso de la IA en materia cautelar, resulta esencial insistir, una vez más, en que ambos presupuestos - *fumus commissi delicti y periculum in mora*- deben valorarse de forma autónoma y sucesiva, de forma que solo se entre a valorar el riesgo de reiteración delictiva, una vez constatada la consistencia del juicio reforzado de imputación. Además, hay que tener en cuenta también que la reiteración delictiva no ha de valorarse en abstracto, sino en relación con el concreto

---

varios Estados de Estados Unidos tanto para presos provisionales como definitivos, con base en datos de presos previos, así como con datos del encausado o preso en cuestión, que ha de responder un cuestionario con más de 100 ítems distintos. Las críticas al sistema COMPAS ponen en entredicho su eficacia. En este sentido, 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). Igualmente, ha sido puesta en entredicho su neutralidad, con acusaciones serias de producir sesgos racistas, puestos de manifiesto, entre otros, por el estudio desarrollado por DIETRICH, William; MENDOZA, Christina; BRENNAN, Tim. COMPAS Risk Scales: Demonstrating accuracy equity and predictive parity. Northpointe Inc. Research Department, 2016, , el cual demostró que el sistema, a pesar de no incluir la raza directamente como parámetro, presenta un sesgo racial, si se analizan los falsos positivos y los falsos negativos. Concretamente, de los negros catalogados con riesgo algo, 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 delinquiron fueron un 47.7%, frente a un 28% de los negros. El sistema es, asimismo, acusado de opacidad, quedando el algoritmo protegido por el derecho al secreto comercial de sus creadores, lo que impide 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 mismos. A pesar de todas estas carencias, el uso de COMPAS para valorar el riesgo de reincidencia fue avalado por la Corte Suprema de Wisconsin en el ya referido caso Wisconsin v. Loomis (Sentencia 13 de julio de 2016: State v. Loomis, 881, N.W.2d 749, 7532 (Wis, 2016)). En este caso, la Corte avaló la legitimidad del uso de COMPAS por unanimidad, siempre que tal instrumento no fuese el único elemento considerado en la sentencia condenatoria, sino que fuese usado únicamente como complemento a la valoración judicial.

delito imputado al encausado<sup>51</sup>. Y, por supuesto, se debe advertir que el resultado que arroje el algoritmo ha de ser valorado por el juez, junto con el resto de datos que obren en la causa y todas las circunstancias concurrentes en la misma, bajo el prisma de excepcionalidad y proporcionalidad que ha de informar, en todo caso, el uso de la prisión provisional.

Por último, hay que aludir a la evitación del riesgo de destrucción de pruebas como otra de las finalidades justificativas de la prisión provisional (art. 503.3º b) LECrim)<sup>52</sup>. Tal y como indica la propia LECrim, *“para valorar la existencia de este peligro se atenderá a la capacidad del investigado o encausado para acceder por sí o a través de terceros a las fuentes de prueba o para influir sobre otros investigados o encausados, testigos o peritos o quienes pudieran serlo”* (art. 503.3º b) III).

Para valorar este riesgo nuevamente puede resultar esencial la combinación de los datos del caso concreto – tipo de delito, existencia de estructura organizativa o de personas afines al encausado con posibilidad de disponer de fuentes de pruebas o alterarlas, tipo de pruebas que se manejan, tipo de delito, carácter más o menos reciente de los hechos, comportamiento cooperativo u obstruccionista del encausado en la investigación, etc.- con datos análogos de casos pasados, para poder estimar estadísticamente el nivel de riesgo, que siempre deberá de ser finalmente interpretado por el juzgador, considerando aspectos como la actitud, la aptitud y la capacidad real del sujeto para materializar objetivamente los riesgos detectados.

---

<sup>51</sup> En este sentido, concuerdo con NIEVA FENOLL, Jordi, op. cit., pp. 74-75, cuando indica que *“lo que es esencial en materia judicial es relacionar el riesgo de reincidencia con el delito precisamente investigado, cuya constancia debe de determinarse al margen del pronóstico de reincidencia”*. De lo contrario se correría el riesgo, en absoluto desdeñable, de crear una tipología de sujetos elegibles para padecer la prisión preventiva, a la vista de datos desvinculados de la autoría delictiva. Y es que, tal y como advierte el autor citado, cuando la herramienta se nutre de datos desvinculados de la autoría delictiva, se convierte en prejuiciosa y racista, tendiendo, poco a poco, a construir *“un modelo único de sociedad perfecta en la que será un potencial delincuente quien no cumpla con esos patrones externos”*.

<sup>52</sup> Concretamente, la medida puede adoptarse para *“evitar la ocultación, alteración o destrucción de las fuentes de prueba relevantes para el enjuiciamiento en los casos en que exista un peligro fundado y concreto”* en este sentido; aunque, en este caso, por un tiempo limitado de 6 meses improrrogables, mientras se aseguran las referidas fuentes de prueba.

Nuevamente, se debe llamar la atención sobre la necesidad de que el juzgador interprete los resultados aportados por la herramienta de IA a la luz de las concretas circunstancias del caso. Y es que, si bien la aptitud para manipular las pruebas quizás pueda ser estimada de forma más o menos fiable por la herramienta, la actitud del investigado en el caso concreto, puede requerir de una interpretación histórica, contextual e incluso emocional de las circunstancias del caso, que difícilmente podrá ser reproducida por un sistema de IA.

Por último, hay que aludir a la variabilidad de las decisiones cautelares, que no sólo deben de poder ser recurridas, sino también revisadas ante cambios en las circunstancias que motivaron su adopción. En este sentido, los sistemas de IA, como colectores de datos, pueden servir para mantener actualizado el nivel de riesgo, o incluso como sistema de alerta ante cambios en las circunstancias iniciales que motivaron la decisión cautelar, siempre que los datos de que se nutren se mantengan actualizados y el algoritmo sea usado con esta funcionalidad por los órganos judiciales: esto es, para reconsiderar decisiones cautelares ya adoptadas ante la aparición de nuevos datos. Por el contrario, si se usan los sistemas de IA para valorar riesgos de manera estática, esto es, como foto fija del nivel de riesgo en un momento dado, existe el peligro de que esos resultados permanezcan inmutables a pesar de producirse cambios relevantes en las circunstancias. El dinamismo que, en este sentido, podría introducirse con un uso adecuado de la IA en el ámbito cautelar es sin duda otra de sus potencialidades.

### **3.- RIESGOS PARA LOS DERECHOS FUNDAMENTALES PROCESALES Y GARANTÍAS A ADOPTAR EN EL USO DE LA IA EN LA TOMA DE DECISIONES CAUTELARES**

La IA, cuando se utiliza para la toma de decisiones en el ámbito de la administración de justicia, plantea riesgos nada desdeñables desde el punto de vista de los derechos fundamentales y, específicamente, de los derechos procesales<sup>53</sup>. Es por ello que para introducir el uso de

---

<sup>53</sup> Sin ánimo de exhaustividad, cabe señalar que el uso de la IA como elemento de apoyo a las decisiones judiciales podría afectar a derechos sustantivos

sistemas de IA, sin restringir indebidamente los derechos y libertades del ciudadano en relación con el debido proceso, se hace necesario observar ciertas garantías en su aplicación<sup>54</sup>.

Para poder hacer uso de las potencialidades de la IA en el ámbito de la justicia, y señaladamente en el contexto de la justicia cautelar penal, es preciso detectar o anticipar los riesgos que su uso puede conllevar, para diseñar estrategias que permitan neutralizar o minimizar tales riesgos. En este sentido, la adopción de una medida restrictiva de derechos fundamentales de primer orden –como es la prisión provisional– ha de venir flanqueada por las debidas garantías para asegurar, en primer lugar,

---

tales como la intimidad, la protección de datos personales o la igualdad ante la ley, así como a diversos derechos procesales, entre los que cabe destacar el derecho a la tutela judicial efectiva, especialmente en lo relativo al ejercicio efectivo del derecho a los recursos, y el derecho de defensa, señaladamente a través de limitaciones al esencial principio de contradicción. Además, en el marco del proceso penal, y específicamente de la justicia cautelar, se encontrarían también en juego derechos fundamentales de primer orden como la presunción de inocencia del encausado o su libertad personal, 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 condena.

<sup>54</sup> Conuerdo, en este punto, con MARTÍN DIZ, Fernando. Aplicaciones de inteligencia artificial en procesos penales por delitos relacionados con la corrupción. En: RODRÍGUEZ-GARCÍA, Nicolás; CARIIZO GONZÁLEZ-CASTELL, Adán; RODRÍGUEZ-LÓPEZ, Fernando (eds.). *Corrupción: compliance, represión y recuperación de activos*, Valencia: Tirant Lo Blanch, pp. 533-568, p. 538, cuando indica que la aplicación de la IA al ámbito del Derecho Procesal ha de atender preferentemente a asegurar las garantías procesales por delante de la eficiencia. De acuerdo con el autor, de nada sirven los innegables avances en este ámbito, si no vienen acompañados del necesario “*respeto al marco de garantías constitucionales fundamentales del justiciable y del funcionamiento de la propia justicia*”. En una línea similar, BUENO DE MATA, Federico. Macrodatos, inteligencia artificial y proceso: luces y sombras. *Revista General de Derecho Procesal*, n. 51, 2020, p. 18, sostiene que la IA como elemento predictivo en juicio y, en general, como elemento de auxilio al juzgador en la toma de decisiones “*debe ser introducida en nuestro sistema con todas las cautelas posibles para respetar todos y cada uno de los principios y las garantías procesales*”. En este sentido, tanto la Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno, como el Libro Blanco sobre la Inteligencia artificial – un enfoque europeo orientado a la excelencia y a la confianza, antes citados, inciden en algunas de las limitaciones a su uso o garantías a observar de cara a su implementación en la administración de justicia.

la legalidad y previsibilidad de la decisión; en segundo lugar, el respeto del derecho a la libertad de quien, no olvidemos, ha de ser considerado y tratado como inocente; en tercer lugar, para garantizar el esencial derecho de defensa del justiciable, asegurando la debida contradicción y las posibilidades de impugnar de forma efectiva la decisión cautelar; y, en cuarto lugar, para 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, pero atendiendo a las concretas y específicas circunstancias del caso concreto.

Así pues, la primera garantía a respetar tiene que ver con el principio de legalidad, el cual exige que exista una previsión normativa del ámbito de aplicación y de la forma de utilización de los sistemas de IA en la justicia<sup>55</sup>, lo que cobra especial relevancia cuando estas herramientas se emplean para tomar decisiones restrictivas de derechos fundamentales, como las relativas a medidas cautelares personales. En este sentido, la ley debiera de especificar, al menos, para qué tipo de riesgos cautelares podrían ser empleados los sistemas de IA, con qué límites y con qué garantías de transparencia, contradicción y control judicial, así como la forma en que sus resultados debieran de reflejarse en la motivación de la resolución judicial<sup>56</sup>.

Pero la previsión legal de su uso siendo necesaria, no es suficiente. El principal escollo que presenta el uso de la IA como instrumento de apoyo al ejercicio de la función jurisdiccional tiene que ver con las limitaciones que este puede implicar desde el punto de vista del derecho de defensa, por las restricciones que podría conllevar para el principio de contradicción, y para el derecho a la tutela judicial efectiva, especialmente en la faceta relativa al efectivo ejercicio del derecho a los recursos<sup>57</sup>. En este sentido,

---

<sup>55</sup> DE HOYOS SANCHO, Montserrat, op. cit., p. 32

<sup>56</sup> En este sentido, DE HOYOS SANCHO, Montserrat, op. cit., p. 33, sostiene que, además de la referencia legal expresa al uso de herramientas de IA en el ejercicio de la función jurisdiccional, debería de existir una regulación suficiente en otros textos normativos, fuera de la LECrim, que permitieran su modificación y actualización de forma flexible, de forma que la legislación no supusiera un freno para los avances técnicos en la materia.

<sup>57</sup> En palabras de ZAVRŠNIK, Aleš, op. cit., p. 14: *“Derivative procedural rights, such as the right to cross-examine witnesses, should be interpreted so as to also encompass the right to examine the underlying rules of the risk-scoring*



la falta de transparencia en cuanto al funcionamiento de los algoritmos, a los datos de los que estos se nutren, al peso ponderado de tales datos en el resultado del algoritmo, así como al peso relativo de la herramienta en la decisión judicial son factores que podrían vulnerar derechamente las posibilidades del encausado de defenderse de forma efectiva y de discutir las decisiones judiciales<sup>58</sup>. Esta situación podría agravarse en fase cautelar, debido a la limitada información manejada por el juzgador, que le puede animar a tomar decisiones más intuitivas o a dejarse llevar por los resultados de la herramienta ante el vacío probatorio imperante<sup>59</sup>.

Para contrarrestar los referidos riesgos al derecho de defensa y a la tutela judicial efectiva resulta absolutamente imprescindible, primeramente, que exista un cierto nivel de transparencia en cuanto al funcionamiento del algoritmo y los datos de que este se nutre. En este sentido, concuerdo con DE HOYOS SANCHO<sup>60</sup> en que no se debería autorizar el uso de herramientas que puedan ser determinantes del sentido

---

*methodology. In probation procedures, this right should entail ensuring that a convicted person has the possibility to question the modelling applied – from the data fed into the algorithm to the overall model design”.*

<sup>58</sup> Según MARTÍN DIZ, Fernando, op. cit., pp. 539 y 544, “la transparencia es exigencia imprescindible en la aplicación de tecnologías de IA que afecten a derechos fundamentales”. En este sentido, el autor apuesta por una “transparencia técnica total”, abogando por un sistema que debe de ser “comprensible con un lenguaje claro y accesible, que describa como se producen los resultados”.

<sup>59</sup> ZAVRŠNIK, Aleš, op. cit., p. 13; GORDON, Faith, op. cit., pp. 162-164; GASCÓN INCHAUSTI, Fernando, op. cit., pp. 204-205; DE HOYOS SANCHO, Montserrat, op. cit., p. 36.

<sup>60</sup> DE HOYOS SANCHO, Montserrat, op. cit., p. 37. Tal y como indica la autora, en este ámbito no pueden admitirse las denominadas “black box”, siendo 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; De la misma opinión es MARTÍNEZ GARAY, Lucía, op. cit., pp. 497-499, quien se muestra muy crítica con la Sentencia dictada en el caso Wisconsin v. Loomis en este punto, ya que la misma vino a legitimar la falta de transparencia del algoritmo, con base en el derecho al secreto empresarial de la entidad que desarrolló COMPAS. De acuerdo con la citada autora, “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”.

de una resolución judicial si las partes no pueden conocer los elementos que integrarán el algoritmo y el propio funcionamiento del sistema de IA. Por lo tanto, en este ámbito no tiene cabida el derecho al secreto industrial si entra en contradicción con el derecho a un proceso justo, al cual los principios de contradicción e igualdad de armas resultan inherentes. Pero, además, tampoco tendrán cabida en este contexto, al menos sin los adecuados contrapesos, aquellos sistemas de aprendizaje automático (*machine learning*) que, por su propio funcionamiento, generan irremediamente cajas negras, sin que sea posible conocer el funcionamiento interno del algoritmo una vez que comienza a arrojar resultados y a generar relaciones autónomas respecto de las predeterminaciones de su creador<sup>61</sup>.

Otra garantía trascendental que debe acompañar a la introducción de la IA en la justicia penal en general, y en la cautelar en particular, tiene que ver con la necesidad de que el sistema sea usado únicamente como complemento o apoyo a las decisiones judiciales, implementándose sistemas *human-in-command* (HIC), caracterizados por la necesaria intervención humana en la decisión final<sup>62</sup>. Hay quien se refiere, incluso, a un derecho

---

<sup>61</sup> Sobre las dificultades de garantizar la transparencia en sistemas de aprendizaje automático, véase VEALE, Michael; EDWARDS, Lilian. Clarity, surprises, and further questions in the Article 29. Working Party draft guidance on automated decision-making and profiling. *Computer Law & Security Review*, n. 34, pp. 398–404, 2018.

<sup>62</sup> SIMÓN CASTELLANO, Pere, op. cit., pp. 144, 214 y 215. El autor sostiene que la principal garantía en la aplicación de sistemas de valoración de riesgos en la justicia pasa, precisamente, por los jueces, que no deben de aceptar decisiones automatizadas, siendo esencial que ponderen los resultados de la herramienta y exterioricen en su motivación en qué medida han sido atendidos o no. En la misma línea, sostiene DE HOYOS SANCHO, Montserrat, op. cit., p. 35, que la decisión final no puede basarse exclusivamente en lo que resulte de la aplicación del algoritmo; del mismo modo que no sería aceptable que el juez delegase en un perito la resolución sobre la pretensión punitiva, por muy complejo que sea el objeto de la pericia. Este argumento, a mi juicio, resulta enteramente trasladable a la decisión cautelar; Por su parte, MARTÍN DIZ, Fernando, op. cit., pp. 536-537, también alude a la función de mero auxilio o asistencia que, en el estadio actual de desarrollo, puede desempeñar la IA como elemento en la toma de decisiones jurisdiccionales; Igualmente, según BUENO DE MATA, Federico, op. cit., pp. 27-29, hoy en día, la IA no puede ser usada como un elemento sustitutivo de los operadores jurídicos, sino únicamente como elemento de auxilio, poniéndose en riesgo, en otro caso, el propio concepto de función jurisdiccional en tanto que actividad atribuida

de nuevo cuño en este ámbito, consistente en que las decisiones no se basen exclusivamente en el tratamiento automatizado de datos<sup>63</sup>. Esta garantía tiene que ver con la característica de jurisdiccionalidad que informa las medidas cautelares, y señaladamente la prisión provisional, la cual sólo puede ser acordada por una autoridad judicial<sup>64</sup>.

Pues bien, vinculado con lo anterior, y precisamente como forma de controlar la influencia del juicio humano y del juicio algorítmico sobre las decisiones judiciales, resulta absolutamente esencial insistir en la necesidad de motivar exhaustivamente las decisiones sobre tutela cautelar y el peso relativo que, en su caso, hayan tenido en las mismas los resultados del algoritmo, junto con el resto de factores tomados en consideración<sup>65</sup>. Solo de este modo es posible garantizar el efectivo derecho a los recursos<sup>66</sup>, contenido básico del derecho fundamental a la

---

de forma exclusiva y excluyente a los jueces y magistrados integrantes del Poder Judicial.

<sup>63</sup> SIGNORATO, Silvia. Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo. *Rivista di diritto processuale*, v. 75, n. 2, pp. 605-616, p. 613, 2020.

<sup>64</sup> Según el TC “[la] prisión provisional es una medida cautelar que sólo puede ser acordada por los órganos judiciales [...] desde la perspectiva de que toda medida restrictiva de derechos fundamentales requiere una decisión judicial motivada” (STC 147/2000, de 29 de mayo, FJ 4 b))

<sup>65</sup> Es evidente que la motivación de las resoluciones judiciales es presupuesto indispensable para la efectividad del derecho a los recursos. En este sentido, SIMÓN CASTELLANO, Pere, op. cit., p. 178, sostiene que la necesidad de garantizar la capacidad del investigado de impugnar y recurrir las decisiones judiciales, cuando estas se apoyan en sistemas de IA, implica la capacidad de impugnar tanto la decisión en sí como el funcionamiento del sistema de IA, para lo cual es preciso que el encausado conozca el peso que se ha dado a cada factor, cual es la fórmula algorítmica, cómo se obtienen esos resultados y qué influye más o menos en la decisión final. En una línea similar, NIEVA FENOLL, Jordi, op. cit., p. 73, advierte de la dificultad de defenderse de un algoritmo sin conocer su contenido, así como de la imposibilidad de que los tribunales se escuden en la dificultad técnica de conocer el funcionamiento interno del algoritmo, especialmente cuando lo usan para tomar decisiones que afectan a la libertad del justiciable, todo lo cual se traduce en la necesidad de hacer transparente el uso del algoritmo a través de la imprescindible motivación judicial.

<sup>66</sup> En este sentido, el *Libro Blanco sobre la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza* (p. 29) indica que, “en el caso de las aplicaciones de IA de riesgo elevado y de otras aplicaciones de IA, debe

tutela judicial efectiva<sup>67</sup>, así como el derecho de defensa y el principio de contradicción consustanciales a la justicia penal. En el caso de la prisión provisional, dada su excepcionalidad y carácter especialmente restrictivo y grave respecto de los derechos del encausado, la motivación cobra una relevancia excepcional, debiendo de realizar el juzgador un esfuerzo motivacional extra que justifique la necesidad y proporcionalidad de la medida en atención a las circunstancias del caso concreto y a las circunstancias individuales del investigado<sup>68</sup>.

Por último, destacar que otro de los riesgos más relevantes que se derivan del uso de sistemas de IA en la fase cautelar es el vinculado a posibles vulneraciones del derecho a la libertad individual y a la presunción de inocencia ante un tratamiento automatizado de datos masivos que sirva para decidir sobre la libertad personal de un individuo sin valorar suficientemente sus circunstancias personales. En este sentido, cabe recordar que el riesgo de reiteración delictiva ha de valorarse en atención a criterios objetivos presentes en el caso concreto, así como en relación con el delito presuntamente cometido, sin que sea suficiente con una valoración de la predisposición a delinquir general del sujeto ni, mucho menos, de una tipología de individuo<sup>69</sup>. De nuevo, como cautela para prevenir estos riesgos de sobre-generalización de ciertas características del encausado o de la causa como elemento predictivo de riesgo, la motivación es clave, debiendo exteriorizarse el análisis particularizado del caso por parte del juzgador. La pérdida de eficacia y rapidez en la toma de decisiones que puede suponer este ejercicio de valoración casuística del asunto, sin duda se

---

*garantizarse una acción judicial efectiva para las partes que hayan sufrido repercusiones negativas derivadas de los sistemas de IA”.*

<sup>67</sup> Ver, por todas, STC 96/2021, de 10 de mayo de 2021.

<sup>68</sup> En este punto, se hace necesario llamar la atención sobre el riesgo de que la estandarización o automatización de la actividad jurisdiccional conduzca a motivaciones “modelo” que no respeten las exigencias del derecho a la tutela judicial efectiva (Véase, sobre esta idea, MARTÍN DIZ, Fernando, op. cit., p. 565).

<sup>69</sup> Una vez más, cabe referirse, en este punto, a los argumentos de la Corte Suprema de Wisconsin en el tantas veces citado caso Loomis, en el cual se indicó que los jueces, cuando manejan herramientas de predicción de riesgos como COMPAS, deben de aplicar a sus resultados un margen de discrecionalidad, poniéndolos en el contexto del resto de circunstancias particulares propias del caso y del individuo concretos.

ve compensado por las mayores garantías que tal proceder implica para el justiciable, evitando así incurrir en discriminaciones frente a determinados individuos o colectivos<sup>70</sup>. También en este punto resulta relevante hacer hincapié, una vez más, en el carácter meramente auxiliar o asistencial que deben de tener los sistemas de IA en la toma de decisiones que afecten a derechos fundamentales, de tal forma que la valoración humana sirva para aplicar determinados controles y ponderaciones que aseguren un uso ético y no sesgado del resultado algorítmico<sup>71</sup>.

## CONCLUSIONES

El uso de sistemas de IA en la toma de decisiones cautelares presenta enormes potencialidades, en tanto que permite realizar valoraciones de los riesgos que integran el *periculum in mora* fundamentadas en el manejo de muchos más datos que aquellos que la memoria y/o experiencia del juzgador puede gestionar. En principio, a mayor cantidad de datos, mejores decisiones serán ofrecidas por el sistema, a condición de que los datos recogidos tengan una calidad suficiente. En este sentido, respondiendo al principal interrogante planteado en este trabajo, cabe afirmar que el uso de la IA puede mejorar la eficacia del proceso de toma de decisiones en el ámbito de la justicia cautelar y, concretamente, puede servir de elemento de auxilio relevante en lo que se refiere a la valoración de los riesgos o peligros que justifican la adopción de la prisión provisional.

Sin embargo, al igual que el razonamiento humano, los algoritmos no están exentos de sesgos y riesgos de discriminación. Algunos de los sesgos presentes en el razonamiento judicial, a través de los denominados

---

<sup>70</sup> En este sentido, tal y como indica MARTÍN DIZ, Fernando, op. cit., p. 542, “*la IA no puede estar diseñada, ni ser aplicada, generando discriminaciones, creación de perfiles con sesgo, marginación o exclusión de colectivos especialmente vulnerables*”. Es por ello que el tratamiento de aquellos datos que puedan aportar información sobre aspectos tales como “*orígenes raciales, genéticos o étnicos, datos biométricos o sanitarios, posiciones socioeconómicas, opiniones políticas o religiosas, pertenencia a determinados sindicatos o grupos u orientación sexual*” resulta especialmente sensible y debe manejarse con elementos correctores que permitan limitar o neutralizar el riesgo de discriminación.

<sup>71</sup> BUENO DE MATA, Federico, op. cit., pp. 28-29.

heurísticos -o atajos del pensamiento humano-, se encuentran presentes en el funcionamiento de los sistemas de IA y, especialmente, en los datos de los que tales sistemas se nutren.

Las correlaciones que se establecen entre determinados parámetros o caracteres desvinculados de la comisión de delitos, pero que terminan por vincularse con el nivel de riesgo de reincidencia o con el riesgo de fuga, junto con la opacidad en el funcionamiento de los algoritmos, pueden terminan por producir resultados injustos y difícilmente controvertibles por quienes los sufren. A esto se suma la inercia del fiscal y el juzgador que, en un contexto de incertidumbre y ausencia de elementos de convicción suficientes, podrían tender a dejarse llevar por los resultados del sistema algorítmico, incluso acomodando su razonamiento y motivación a los mismos<sup>72</sup>.

Ante la situación descrita, es necesario adoptar ciertas cautelas o garantías que aseguren una introducción de la IA en el sistema procesal penal respetuosa con los derechos de los investigados, partiendo de la base de que toda introducción de nuevas tecnologías en el ámbito de la actividad procesal que busque mejoras en su eficacia ha de pasar necesariamente por el debido respeto a los derechos fundamentales en juego<sup>73</sup>. En este sentido, y como principales salvaguardas a adoptar, cabe destacar la importancia de que el uso de la IA en el sistema de justicia penal sea previsto a nivel legal con cierta precisión; que se asegure la transparencia en cuanto al funcionamiento de los algoritmos, así como a los datos de los que estos se nutren; que se exteriorice en la motivación judicial el uso de estos instrumentos, así como su peso relativo en la decisión final del juzgador, de tal forma que se garantice el derecho del justiciable a un recurso efectivo; que el resultado algorítmico sea empleado solo como complemento a las decisiones judiciales, sin llegar a sustituir, en ningún caso, el juicio jurisdiccional; y que, precisamente, ese control

---

<sup>72</sup> Incluso en el estadio actual de sistemas no vinculantes y semiautomáticos de toma de decisiones, el proceso de alcanzar una decisión se modifica. Se modifica también la percepción de responsabilidad por la decisión final. Y el decisor se ve inclinado a retorcer sus propias valoraciones para que coincidan con las del modelo (ZAVRŠNIK, Aleš, op. cit., p. 13). Sobre esta idea, tan inquietante como sugerente, véase GORDON, Faith, op. cit., pp. 162-164.

<sup>73</sup> MARTÍN DIZ, Fernando, op. cit., p. 538.

judicial sirva para relativizar el peso del algoritmo en una decisión que ha de ser necesariamente humana, en tanto humano es el problema que encierra el proceso<sup>74</sup>.

Además de tratar de implementar cautelas que contrarresten los referidos sesgos – ya en la construcción del algoritmo y en la recogida de datos, ya en la interpretación que se haga de los resultados que el sistema arroje– hay que tener presente la limitación de la IA como herramienta apta para la reconstrucción de hechos y la valoración casuística del *fumus commisi delicti*, juicio de imputación reforzado, esencial para la adopción de toda medida cautelar, que no ha de verse contaminado por los resultados entorno a los riesgos de reiteración delictiva, al perfil criminológico del acusado o a la apreciación del riesgo de fuga. Por lo tanto, una vez más, cabe hacer hincapié en el uso limitado que la IA puede tener en las decisiones cautelares de prisión provisional.

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 tales decisiones 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 último término, 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.

---

<sup>74</sup> En este punto resulta interesante recordar las palabras de VÁZQUEZ SO-TELO, José Luis. *Presunción de inocencia del imputado e íntima convicción del tribunal. Estudio sobre la utilización del imputado como fuente de prueba en el proceso penal español*. Barcelona: Bosch, 1985, pp. 470-479, esp. 478-479, al señalar que “*si al perito toca resolver un problema científico, o una duda técnica, al juez corresponde resolver el problema del proceso, que es siempre un problema humano y de proyección social*”. Pues bien, llevando este razonamiento al ámbito de esta investigación, cabría señalar que, si al algoritmo corresponde resolver un problema estadístico, al juez le corresponde interpretar y aplicar ese resultado a la resolución del asunto sometido a su consideración, que, efectivamente, siempre será un problema humano y de proyección social.

## REFERENCIAS

BALCELLS, Marc. Luces y sombras del uso de la inteligencia artificial en el sistema de justicia penal. En: CERRILLO I MARTÍNEZ, Agustí; PEGUERA POCH, Miquel (coords.). *Retos jurídicos de la inteligencia artificial*, Cizur Menor (Navarra): Aranzadi, 2020, pp. 145-159.

BANACH-GUTIERREZ, Joanna Beata. Globalized Criminal Justice in the European Union Context. How Theory Meets Practice. *New Journal of European Criminal Law*, v. 4, n. 1-2, 2013. <https://doi.org/10.1177/203228441300400110>

BORGES BÁZQUEZ, Raquel. El sesgo de la máquina en la toma de decisiones en el proceso penal. *Ius et scientia*, v. 6, n. 2, pp. 54-71, 2020. <https://doi.org/10.12795/ietscientia.2020.i02.05>

BORRÀS ANDRÉS, Núria. La verdad y la ficción de la inteligencia artificial en el proceso penal. En: CONDE FUENTES, Jesús; SERRANO HOYO, Gregorio (dirs.). *La justicia digital en España y la Unión Europea: Situación actual y perspectivas de futuro*. Barcelona: Atelier, 2019, pp. 31-39.

BUENO DE MATA, Federico. Macrodatos, inteligencia artificial y proceso: luces y sombras. *Revista General de Derecho Procesal*, n. 51, 2020.

CHAPMAN, Loren James. Illusory correlation in observational report. *Journal of Verbal Learning & Verbal Behavior*, v. 6, n. 1, pp. 151-155, 1967. [https://doi.org/10.1016/S0022-5371\(67\)80066-5](https://doi.org/10.1016/S0022-5371(67)80066-5)

DE HOYOS SANCHO, Montserrat. El Libro Blanco sobre inteligencia artificial de la Comisión Europea: reflexiones desde las garantías esenciales del proceso penal como 'sector del riesgo', *Revista Española de Derecho Europeo*, n. 76, pp. 9-44, 2020.

DIETRICH, William; MENDOZA, Christina; BRENNAN, Tim COMPAS Risk Scales: Demonstrating accuracy equity and predictive parity. Northpointe Inc. Research Department, 2016. Disponible en: [https://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica\\_Commentary\\_Final\\_070616.pdf](https://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf). Último acceso: 03/09/2021.

DRESSEL, J; FARID, H. The accuracy, fairness and limits of predicting recidivism. *Science advances*, n. 4, 2018. <https://doi.org/10.1126/sciadv.aao5580>

ESTÉVEZ MENDOZA, Lucana. Algoritmos policiales basados en IA derechos fundamentales a la luz de Hart y Valcri. En JIMÉNEZ CONDE, Fernando; BELLIDO PENADÉS, Rafael (coords.). *Justicia: garantías vs. Eficacia*. Valencia: Tirant Lo Blanch, 2019, pp. 665-674.



GARCÍA ESPAÑA, Elisa. Extranjeros sospechosos, condenados y ex condenados: Un mosaico de exclusión. *Revista Electrónica de Ciencia Penal y Criminología*, n. 19-15, 2017. Disponible en: <http://criminnet.ugr.es/recpc/19/recpc19-15.pdf>. Último acceso: 03/09/2021.

GASCÓN INCHAUSTI, Fernando. Desafíos para el proceso penal en la era digital: externalización, sumisión pericial e inteligencia artificial. En CONDE FUENTES, Jesús; SERRANO HOYO, Gregorio (dirs.). *La justicia digital en España y la Unión Europea: Situación actual y perspectivas de futuro*. Barcelona: Atelier, 2019, pp. 191-206.

GIALUZ, Mitja. Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa. *Diritto Penale Contemporaneo*, 2019, pp. 19-20. Disponible en: <https://archiviodpc.dirittopenaleuomo.org/upload/6903-gialuz2019b.pdf>. Último acceso: 03/09/2021.

GIMENO SENDRA, Vicente. *Manual de derecho procesal penal*. 2ª ed. Madrid: Ediciones Jurídicas Castillo de Luna, 2018.

GÓMEZ ORBANEJA, Eduardo. *Comentarios a la Ley de Enjuiciamiento Criminal*, Tomo I, Barcelona: Bosch, 1947.

GORDON, Faith. Book Review on EUBANKS, Virginia. Automating inequality: How high-tech tools profile, police, and punish the poor, *Law, Technology and Humans*, St Martin's Press, New York, v. 1, pp. 162-164, 2018. <https://doi.org/10.5204/lthj.v1i0.1386>

GUERRA PÉREZ, Cristina. *La decisión judicial de prisión preventiva*, Valencia: Tirant Lo Blanch, 2010.

HERNÁNDEZ GIMÉNEZ, María. Inteligencia artificial y Derecho Penal. *Actualidad Jurídica Iberoamericana*, n. 10 bis, pp. 792-843, junio 2019.

JENNINGS, Dennis L.; AMABILE, Teresa M.; ROSS, Lee. Informal covariation assessment: Data-based versus theory-based judgment. En: KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982, pp. 211-230. <https://doi.org/10.1017/CBO9780511809477.016>

KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, 1982. <https://doi.org/10.1017/CBO9780511809477>

KAHNEMAN, Daniel; TVERSKY, Amos, "Judgments of and by representativeness". En: KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under*

*Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982, pp. 84-98. <https://doi.org/10.1017/CBO9780511809477.007>

KAHNEMAN, Daniel; TVERSKY, Amos. Availability: A heuristic for judging frequency and probability. En KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982, pp. 163-178. <https://doi.org/10.1017/CBO9780511809477.012>

KEHL, Danielle; GUO, Priscilla; KESSLER, Samuel. Algorithms in the Criminal Justice System: assessing the use of risk assessment in sentencing, Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School. Disponible en: <https://dash.harvard.edu/handle/1/33746041>. Último acceso: 03/09/2021.

KLEINBERG, Jon; LAKKARAJU, Himabindu; LESKOVEC, Jure; LUDWIG, Jens; MULLAINATHAN, Sendhil. Human Decisions and Machine Predictions. *The Quarterly Journal of Economics*, v. 133, n. 1, pp. 237-293, 2018. <https://doi.org/10.3386/w23180>

LJUNGQUIST, Thomas. Mutual Recognition of Non-Custodial Pre-Trial Supervision Measures in the European Union. *Revue internationale de droit pénal*, v. 77, n. 1-2, 2006. <https://doi.org/10.3917/ridp.771.0169>

MARTÍN DIZ, Fernando. Aplicaciones de inteligencia artificial en procesos penales por delitos relacionados con la corrupción. En: RODRÍGUEZ-GARCÍA, Nicolás; CARIIZO GONZÁLEZ-CASTELL, Adán; RODRÍGUEZ-LÓPEZ, Fernando (eds.). *Corrupción: compliance, represión y recuperación de activos*, Valencia: Tirant Lo Blanch, pp. 533-568.

MARTÍNEZ GARAY Lucía. Peligrosidad, Algoritmos y *Due Process*: el Caso *State V Loomis*. *Revista de Derecho Penal y Criminología*, 3.ª Época, n. 20, pp. 485-502, julio de 2018. <https://doi.org/10.5944/rdpc.20.2018.26484>

MIRÓ LLINARES, Fernando. Inteligencia artificial y Justicia Penal: Más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, n. 20, pp. 87-130, 2018. <https://doi.org/10.5944/rdpc.20.2018.26446>

MIRÓ LLINARES, Fernando. “El modelo policial que viene: mitos y realidades del impacto de la inteligencia artificial y la ciencia de datos en la prevención policial del crimen”. En MARTÍNEZ ESPASA, José (coord.). *Libro blanco de la prevención y seguridad local valenciana: Conclusiones y propuestas del Congreso Valenciano de Seguridad Local: la prevención del siglo XXI, celebrado en Benidorm del 16 al 18 de noviembre de 2018*, Valencia: Agència Valenciana de Seguretat i Resposta a les Emergències; Agència Valenciana de Seguretat i Resposta a les Emergències,

Instituto Valenciano de Seguridad Pública y Emergencias, 2019, pp. 98-113. Disponible en: <https://dialnet.unirioja.es/servlet/libro?codigo=739767>. Último acceso: 03/09/2021.

MUÑOZ RODRÍGUEZ, Ana Belén. El impacto de la inteligencia artificial en el proceso penal. *Anuario de la Facultad de Derecho. Universidad de Extremadura*, n. 36, pp. 695-728, 2020. Último acceso: 03/09/2021. <https://doi.org/10.17398/2695-7728.36.695>

NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*. Madrid: Marcial Pons, 2018.

O'NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown, 2016.

PÉREZ ESTRADA, Miren Josune. El uso de algoritmos en el proceso penal y el derecho a un proceso con todas las garantías. En: BARONA VILAR, Silvia (ed.) *Claves de la justicia penal: feminización, inteligencia artificial, supranacionalidad y seguridad*. Valencia: Tirant lo Blanch, 2019, pp. 235-154.

RAFARACI, Tomasso. The Application of the Principle of Mutual Recognition to Decisions on Supervision Measures as an Alternative to Provisional Detention. En: RUGGERI, Stefano (ed.). *Liberty and Security in Europe*, Osnabruck: Universitätsverlag, 2013.

RAMOS MÉNDEZ, Francisco. *Enjuiciamiento criminal. Décima lectura constitucional*. Barcelona: Atelier, 2011.

RECIO JUÁREZ, Matías. Nuevos Instrumentos para el Cumplimiento Transnacional de las Medidas Cautelares Alternativas a la Prisión Provisional en la Unión Europea. *Revista Direito e Inovação*, v. 2, n. 2, pp. 198-213, 2014.

ROMEO CASABONA, Carlos María. Inteligencia artificial, derechos fundamentales y proceso penal. *Comunicaciones en propiedad industrial y derecho de la competencia*, n. 89, pp. 253-271, enero-abril de 2020.

SIGNORATO, Silvia. Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo. *Rivista di diritto processuale*, v. 75, n. 2, pp. 605-616, 2020.

SIMON CASTELLANO, Pere. *Justicia cautelar e inteligencia artificial*. La alternativa a los atáxicos heurísticos judiciales Barcelona: Bosch, 2021.

SOLAR CAYÓN, José Ignacio. Inteligencia artificial en la justicia penal: los sistemas algorítmicos de evaluación de riesgos. En: SOLAR CAYÓN, José Ignacio (ed.).

*Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho*, Alcalá de Henares: Universidad de Alcalá de Henares, 2020, pp. 125-172.

VÁZQUEZ SOTELO, José Luis. *Presunción de inocencia del imputado e íntima convicción del tribunal. Estudio sobre la utilización del imputado como fuente de prueba en el proceso penal español*. Barcelona: Bosch, 1985.

VEALE, Michael; EDWARDS, Lilian. Clarity, surprises, and further questions in the Article 29. Working Party draft guidance on automated decision-making and profiling. *Computer Law & Security Review*, n. 34, pp. 398–404, 2018. <https://doi.org/10.1016/j.clsr.2017.12.002>

ZAVRŠNIK, Aleš. Algorithmic justice: Algorithms and big data in criminal justice setting. *European Journal of criminology*, pp. 1-20, pp. 13-14, septiembre de 2019. <https://doi.org/10.1177/1477370819876762>

### **Additional information and author's declarations (scientific integrity)**

*Conflict of interest declaration:* the author confirms that there are no conflicts of interest in conducting this research and writing this article.

*Declaration of authorship:* all and only researchers who comply the authorship requirements of this article are listed as authors; all coauthors are fully responsible for this work in its entirety.

*Declaration of originality:* the author assures that the text here published has not been previously published in any other resource and that future republication will only take place with the express indication of the reference of this original publication; she also attests that there is no third party plagiarism or self-plagiarism.

### Editorial process dates

(<http://www.ibraspp.com.br/revista/index.php/RBDPP/about/editorialPolicies>)

- Submission: 15/07/2021
- Desk review and plagiarism check: 18/07/2021
- Review 1: 27/07/2021
- Review 2: 30/07/2021
- Preliminary editorial decision: 01/09/2021
- Correction round return: 08/09/2021
- Final editorial decision: 01/10/2021

### Editorial team

- Editor-in-chief: 1 (VGV)
- Associated-editor: 2 (BC, LG)
- Reviewers: 2

### HOW TO CITE (ABNT BRAZIL):

NEIRA PENA, Ana María. Inteligencia artificial y tutela cautelar. Especial referencia a la prisión provisional. *Revista Brasileira de Direito Processual Penal*, vol. 7, n. 3, p. 1897-1933, set./dez. 2021. <https://doi.org/10.22197/rbdpp.v7i3.618>



Esta obra está licenciada com uma Licença *Creative Commons Atribuição-NãoComercial 4.0 Internacional*.