



UNIVERSITAT DE
BARCELONA

Facultat de Dret

Inteligencia Artificial y la prueba del
periculum libertatis

Trabajo de Final de Grado

Paula Belén Pérez Chávez

NIUB: 20080535

Derecho Procesal Penal

Tutor: Jordi Nieva Fenoll

Curso: 2020 – 2021

Segundo cuatrimestre

*“Cuando los ordenadores tomen el control,
quizá ya no lo podamos volver a recuperar.
Sobreviviremos mientras ellos nos toleren.”*

Marvin Lee Minsky, 1970

Índice

ABREVIATURAS	1
INTRODUCCIÓN	2
1. LA INCORPORACIÓN DE LA INTELIGENCIA ARTIFICIAL EN EL PROCESO PENAL	4
1.1. CONCEPTUALIZACIÓN.....	4
1.2. NOCIÓN BÁSICA DEL FUNCIONAMIENTO DE LA INTELIGENCIA ARTIFICIAL...	5
1.3. USOS ACTUALES DE LA INTELIGENCIA ARTIFICIAL EN EL PROCESO PENAL...	6
1.3.1. <i>Detección de denuncias falsas</i>	6
1.3.2. <i>Recogida de vestigios en la escena del crimen</i>	7
1.3.3. <i>Análisis de la escena del crimen</i>	7
1.3.4. <i>Reconstrucción de hechos</i>	8
2. EL PROCEDIMIENTO COGNITIVO EN LA VALORACIÓN DEL RIESGO Y LA INTELIGENCIA ARTIFICIAL	9
2.1. SISTEMA EXPERIENCIAL	10
2.1.1. <i>Heurístico de accesibilidad</i>	11
2.1.2. <i>Heurístico de afecto</i>	13
2.2. LA INFLUENCIA DE LAS EMOCIONES EN LA VALORACIÓN DEL RIESGO	13
3. LA PRUEBA DEL <i>PERICULUM LIBERTATIS</i> Y LA INTELIGENCIA ARTIFICIAL	15
3.1. EVALUACIÓN DEL RIESGO Y LA INTELIGENCIA ARTIFICIAL	15
3.1.1. <i>Riesgo de fuga</i>	16
3.1.2. <i>Riesgo de destrucción de pruebas</i>	17
3.1.3. <i>Riesgo de reiteración delictiva</i>	18
3.2. EL PROBLEMA DEL CRITERIO DE PREDICCIÓN.....	25
4. PROBLEMÁTICA DEL USO DE LA INTELIGENCIA ARTIFICIAL EN LA VALORACIÓN DEL RIESGO	27
4.1. EL SESGO DE LOS DATOS	27
4.2. USO ÉTICO DE LA INTELIGENCIA ARTIFICIAL	29
4.3. DERECHOS DEL REO.....	30
4.3.1. <i>Derecho a un juez imparcial</i>	31
4.3.2. <i>Derecho de defensa</i>	33
4.3.3. <i>Derecho a la presunción de inocencia</i>	34
CONCLUSIONES	36
FUENTES Y BIBLIOGRAFÍA.....	39

Abreviaturas

BOE	Boletín Oficial del Estado
CE	Constitución Española
CEDH	Convenio Europeo de Derechos Humanos
DUDH	Declaración Universal de Derechos Humanos
LECrim	Ley Enjuiciamiento Criminal
LOPJ	Ley Orgánica del Poder Judicial
PIDCP	Pacto Internacional de Derechos Civiles y Políticos
TC	Tribunal Constitucional
TEDH	Tribunal Europeo de Derechos Humanos

Introducción

Es innegable el hecho de que nuestro mundo gira alrededor de la inteligencia artificial hasta tal punto que, en ocasiones, es utilizada sin darnos cuenta. La inteligencia artificial está sumamente presente en nuestras vidas, desde la recomendación de una determinada película en *Netflix* según nuestros gustos, hasta la sugerencia de elección de la ruta óptima que nos proporciona nuestro navegador GPS basado en la afluencia de tráfico en un tramo concreto.

La capacidad de la inteligencia artificial de tener en consideración multitud de datos supera con creces la del ser humano, haciéndola así, una herramienta realmente beneficiosa para el mismo. Tal es esta capacidad que la inteligencia artificial “controla” ciertos fenómenos que el ser humano no realiza adecuada o completamente¹.

La inteligencia artificial es también utilizada en el ámbito jurídico de diferentes formas, empezando por la mecanización de servicios judiciales –tal como la búsqueda jurisprudencial– hasta la predicción de resultados legales, siendo este último el más controvertido². Asimismo, encontramos su aplicabilidad, investigación y desarrollo en el proceso penal en cuanto a materia probatoria, detección de denuncias falsas, apoyo en la investigación e incluso en la elaboración de estrategias de acusación y defensa, entre otros³.

El presente trabajo se enmarca en la fase de instrucción del proceso penal, concretamente en la prueba del *periculum libertatis* para la adopción de medidas cautelares. El *periculum libertatis* se concreta en tres posibles riesgos: riesgo de fuga, riesgo de destrucción de pruebas y riesgo de reiteración delictiva. Nos encontramos con la dificultad de que la ley exige el arduo trabajo a los jueces de estimar la concurrencia de estos riesgos en un breve período de tiempo, lo que muchas veces conlleva a la falta de motivación en los autos que acuerdan las medidas cautelares⁴. En consecuencia, y atendiendo a los beneficios que nos proporciona la inteligencia artificial, a lo largo de este trabajo estudiaremos si podemos confiar en su aplicación para la valoración del

¹ PUELLES, Luis, SÁNCHEZ, José y ALBERTOS Pedro, “Inteligencia Artificial e Inteligencia Humana”, *Sociedad Económica de Amigos del País*, 2010, p. 73

² LEGG, Michael. y BELL, Felicity. “Artificial Intelligence and the Legal Profession: A Primer”, *Flip stream*, 2020, p. 13

³ HERNÁNDEZ, María. “Inteligencia artificial y Derecho Penal”, *Actualidad Jurídica Iberoamericana*, 10 bis, 2019, p. 810 y ss.

⁴ NIEVA FENOLL, Jordi. *Derecho procesal III*. Marcial Pons, Madrid, 2017, p. 258 y ss.

periculum libertatis con suficiente fiabilidad como para que pueda afectar a los derechos y libertades del reo mediante la adopción de medidas cautelares.

En primer lugar, haremos una breve aproximación de lo que es la inteligencia artificial y cómo funciona, así como los usos actuales de la misma dentro del proceso penal. A continuación, estudiaremos el procedimiento cognitivo que realiza el juez para la apreciación del riesgo con el fin de observar de qué forma la inteligencia artificial podría evitar los sesgos y prejuicios que muchas veces están presentes en este tipo de valoraciones. Posteriormente analizaremos la posibilidad de la utilización de la inteligencia artificial para la evaluación de los tres riesgos del *periculum libertatis*, haciendo especial referencia a las herramientas actualmente utilizadas. Finalmente estudiaremos los posibles desafíos que puede implicar la utilización de la inteligencia artificial en la adopción de medidas cautelares.

1. La incorporación de la inteligencia artificial en el proceso penal

1.1. Conceptualización

El concepto de “inteligencia artificial” aparece en la década de los 50, pero la idea de que algo “no-humano” ostente inteligencia, apareció mucho antes en la religión y literatura. El término “artificial” denota algo que no es natural, que es falso, y generalmente es asociado a máquinas. El término “inteligencia”, por otro lado, hace referencia a la capacidad de entender o comprender. En cuanto a la noción de “Inteligencia Artificial”, el Diccionario de la Real Academia Española lo define como la “*disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico*”⁵.

En 1936, Alan Turing propuso la noción de máquina universal en la que describía un único algoritmo informático inteligente. La formalización de la inteligencia artificial se inició en 1950 con el *Test de Turing*, una prueba que determina si una máquina es inteligente. Esta prueba consiste en poner a una inteligencia artificial y a un ser humano frente a preguntas que formula un interrogador. Si este no distingue de quien proviene la respuesta, se considerará que la inteligencia artificial es inteligente –fue en 2014 que, por primera vez, un ordenador logró superar el *Test de Turing*⁶–.

El término “inteligencia artificial” existe desde 1956. Su primera aparición fue en una conferencia celebrada en Dartmouth, Estados Unidos, en la que J. McCarthy, M. Minsky, N. Rochester y C. E. Shannon plantearon una propuesta en la que apareció por primera vez el concepto. Esta conferencia tenía como objeto estudiar el trabajo de Alan Turing sobre aprendizaje automático –*machine learning*–. La definición propuesta por J. McCarthy fue la siguiente: “*El objetivo de la inteligencia artificial es desarrollar máquinas que se comporten como si fueran inteligentes*”⁷.

No existe, sin embargo, una definición consensuada. Una aproximación de lo que es la inteligencia artificial es que “*describe la posibilidad de que las máquinas, en alguna*

⁵ Diccionario de la Real Academia Española.

⁶ Sobre este logro, es interesante la lectura de la noticia de FRESNEDA, Carlos. “Un ordenador logra superar por primera vez el test de Turing”. *El Mundo*, 19 de junio de 2014 (disponible en <https://www.elmundo.es/ciencia/2014/06/09/539589ee268e3e096c8b4584.html>; última consulta 20/02/2021).

⁷ GUTIÉRREZ, Juan David. “Retos éticos de la inteligencia artificial en el proceso judicial”, *XLI Congreso Colombiano de Derecho Procesal*, 2020, cit., p.4

medida “piensen”, o más bien imiten el pensamiento humano a base de aprender y utilizar las generalizaciones que las personas usamos para tomar nuestras decisiones habituales”⁸.

1.2. Noción básica del funcionamiento de la inteligencia artificial

Tal y como hemos explicado anteriormente, la inteligencia artificial simula el pensamiento humano. Nuestro cerebro, al nacer, actúa como un disco duro vacío. Necesitamos años para aprender determinados conceptos, empezando por los términos más básicos como las habilidades psicomotrices –andar, correr, saltar– hasta realizar actividades más complejas como, por ejemplo, aprender a tocar el piano. Siguiendo con el último ejemplo, cuando un niño aprende a tocar el piano pone en práctica lo que le ha enseñado previamente un profesor. Se equivoca varias veces hasta que finalmente se va perfeccionando con el tiempo. La inteligencia artificial “aprende” de forma similar.

Generalmente, cuando hablamos de inteligencia artificial, hablamos de los denominados algoritmos de aprendizaje automático o *machine learning*. Un algoritmo es un conjunto de operaciones matemáticas que se entrelazan entre sí con el fin de hallar la solución a un problema⁹.

Este tipo de algoritmos, tal y como indica su nombre, aprenden de forma autónoma. No es necesario que se le introduzcan normas, sino que aprende analizando ejemplos y estableciendo patrones –necesariamente estadísticos–. Por ejemplo, si queremos que una inteligencia artificial que utiliza estos algoritmos nos sepa distinguir una imagen de un gato de la de un perro, deberemos proporcionarle diferentes imágenes de perros y gatos. Conforme van aumentando el número de ejemplos que se le introducen, los patrones que establecen los algoritmos de aprendizaje automático van mejorando, aprendiendo características distintivas para la clasificación de la imagen de un perro o un gato¹⁰.

No obstante, los algoritmos más recientes de aprendizaje automático implican también una problemática: el proceso que siguen para la toma de la decisión se realiza a partir de muchos datos, y las relaciones que establece resultan incomprensibles para la

⁸ NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018, cit. p. 20.

⁹ BORGES, Raquel. “El sesgo de la máquina en la toma de decisiones en el proceso penal”. *Ius et scientia*, vol. 6, n. 2, 2020, p. 58.

¹⁰ LEGG, Michael. y BELL, Felicity. “Artificial Intelligence and the Legal Profession: A Primer”, *Flip stream*, 2020, p. 7.

mente humana. Una inteligencia artificial nunca nos proporcionará las razones por las cuales ha llegado a una determinada conclusión. Sabemos que establece patrones para aprender a diferenciar, pero no sabemos exactamente en que características se fija¹¹. De aquí que la inteligencia artificial sea entendida como una caja negra –*black box*– en la que entran datos –*inputs*– y salen resultados –*outputs*–.

Actualmente, gracias a las nuevas tecnologías ha habido una gran evolución en la capacidad de procesamiento de datos, dando lugar a nuevas metodologías fundamentadas en la idea de automatización, como el *Deep Learning*, basado en arquitecturas de *Redes Neuronales Artificiales* que se inspiran en la estructura del cerebro humano¹².

1.3. Usos actuales de la inteligencia artificial en el proceso penal

Tradicionalmente, el campo del derecho siempre se ha mantenido alejado del ámbito de la tecnología. Sin embargo, desde hace ya cierto tiempo, nos encontramos con la incorporación de la inteligencia artificial dentro del proceso penal, así como mecanismos que actualmente están en proceso de investigación y desarrollo.

1.3.1. Detección de denuncias falsas

Uno de los fenómenos que implica un gran coste para los cuerpos policiales es el de las denuncias falsas. Por una parte, porque se destinan recursos a la investigación de un delito que no ha sido cometido, y por otra, porque una vez se demuestra la falsedad de la denuncia, es necesaria la investigación del delito de simulación¹³.

En los últimos años se ha ido desarrollando *Veripol*, una herramienta que utiliza inteligencia artificial desarrollada a iniciativa de la Policía Nacional de España por medio de un proyecto con la Universidad Complutense y la Carlos III de Madrid, y La Sapienza de Roma. Esta herramienta tiene por objeto la detección de denuncias falsas interpuestas en casos de robo con violencia e intimidación. Se basa en un algoritmo de aprendizaje automático –*machine learning*– que recoge alrededor de 300 variables y se fija en el texto de la denuncia. Se funda en la idea de que en el momento de poner una denuncia falsa los denunciadores utilizan un patrón similar en su argumentación. Las primeras pruebas se

¹¹ HOLZINGER, Andreas et al. “A glass-box interactive machine learning approach for solving NP-hard problems with the human-in-the-loop”, 2017, p. 2.

¹² MIRÓ, Fernando. “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots”, *Revista de derecho penal y criminología*, vol.3, n. 20, 2018. p. 91.

¹³ GONZÁLEZ-ÁLVAREZ, José Luis; Jorge SANTOS y Miguel CAMACHO, “Policía predictiva en España. Aplicación y retos futuros”. *Fundación Universitaria Behavior & Law*, vol. 6, n.1, 2020, p. 30.

realizaron sobre casos que ya estaban cerrados, obteniendo un 90% de acierto. Actualmente, el acierto es de un 96%¹⁴.

Aunque las decisiones que nos proporciona esta herramienta no son vinculantes ya que un algoritmo no tiene capacidad de decisión, evita que se malgasten recursos policiales y ayuda a tomar decisiones más acertadas desde la apertura de la investigación. Además, podría convertirse en una herramienta disuasoria para aquellos que se planteen presentar una demanda falsa, evitando el delito de simulación¹⁵.

1.3.2. Recogida de vestigios en la escena del crimen

La recogida de vestigios o pruebas materiales del delito van dirigidas a la averiguación de la comisión del hecho delictivo y la participación de los sujetos que hayan podido cometerlo. Por ello, el análisis de la escena del crimen es de vital importancia para la resolución del caso.

Gracias a la inteligencia artificial sería posible predecir en qué lugares se pueden hallar vestigios con mayor probabilidad. La localización de los indicios desde siempre ha dependido de la intuición y experiencia del agente. Esta experiencia, de alguna forma, puede ser sustituida por la inteligencia artificial puesto que puede “aprender” y predecir los lugares donde es más probable que se hallen vestigios basándose en casos similares¹⁶.

1.3.3. Análisis de la escena del crimen

Cuando se investiga un crimen sin resolver, previamente debe compararse con casos anteriores, buscar acontecimientos relacionados con el caso: la hora, *modus operandi* y diferentes detalles de las personas que podrían estar involucradas.

Para facilitar esta tarea policial, se creó el proyecto VACLRI –*Visual Analytics for Sense-making in Criminal Intelligence Analysis*– por la policía de West Midlands – Reino Unido– y de Ambers –Bélgica–. El objetivo de esta herramienta es analizar diferentes elementos de la escena del crimen y compararlos con millones de fuentes de

¹⁴ ÁLVAREZ BARCELONA, Ramón. “La inteligencia artificial de la Policía que desenmascara denuncias falsas”. *La Vanguardia*. 13 de abril de 2019 (disponible en <https://www.lavanguardia.com/tecnologia/20190414/461583468024/veripol-policia-nacional-inteligencia-artificial-algoritmo-denuncias-falsas.html>; última consulta 24/02/2021).

¹⁵ HERNÁNDEZ, María. “Inteligencia artificial y Derecho Penal”, *Actualidad Jurídica Iberoamericana*, 10 bis, 2019, p. 819.

¹⁶ NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*. Marcial Pons, Madrid, 2018, p. 26.

información de la base de datos. Con ello, establece patrones entre los diferentes casos analizados con el fin de reconstruir un “mapa delictivo”¹⁷.

Al funcionar esta herramienta con inteligencia artificial, puede aprender a interpretar el lenguaje humano. Un ejemplo sería el supuesto de que en tres interrogatorios distintos se interrogue a la misma persona. Es posible que el primer y el segundo interrogador utilicen la palabra “tímido” para calificar al individuo, mientras que el tercer interrogador utilice “vergonzoso”. Esta herramienta, al igual que un humano, podrá saber que se está describiendo a la misma persona¹⁸.

1.3.4. Reconstrucción de hechos

En 2006 se inició el desarrollo de *Stevie*, una herramienta que emplea inteligencia artificial capaz de reconstruir hechos a partir de vestigios ya existentes.

Durante la investigación, esta herramienta permite conservar una visión general de toda la información con el fin de poder comparar diferentes escenarios. *Stevie* se basa en casos anteriores, reconstruyendo una hipótesis de lo que podría haber sucedido en el caso en cuestión¹⁹. La historia que proporciona esta herramienta es temporalmente coherente, es decir, descarta historias que sean temporalmente inconsistentes.

Otra herramienta de predicción que utiliza inteligencia artificial es *ECHO*, una aplicación que realiza estrategias de acusación y defensa. Podría ser de especial utilidad tanto para abogados como para la fiscalía. Es utilizada para la preparación de juicios. Permite llevar a cabo una simulación del juicio que se realizará a partir de los vestigios y pruebas provisionales y los argumentos que se espera que se utilicen de la parte contraria²⁰.

¹⁷ ESTÉVEZ, Lucana M. “Los derechos fundamentales ante las nuevas tecnologías: protección o vulneración a la luz de HART y VALCRI”, *Alternativas al sistema de justicia criminal latinoamericano*, p. 61.

¹⁸ REVELL, Timothy. “AI detective analyses police data to learn how to crack cases”. *Newscientist*, 10 de mayo de 2017 (disponible en <https://www.newscientist.com/article/mg23431254-000-ai-detective-analyses-police-data-to-learn-how-to-crack-cases/>; última consulta 01/03/2021).

¹⁹ VAN DEN BRAAK, Susan W. y VREESWIJK, Gerard A. W. “A knowledge representation architecture for the construction of stories based on interpretation and evidence”, 2006, p. 1.

²⁰ NISSAN, Ephraim. “Digital technologies and artificial intelligence’s present and foreseeable impact on lawyering, judging, policing and law enforcement”. *Sprinter*, 2015, p. 451.

2. El procedimiento cognitivo en la valoración del riesgo y la inteligencia artificial

Para la evaluación de los tres riesgos que forman el *periculum libertatis*, el juez lleva a cabo un procedimiento cognitivo que muchas veces implica que esta evaluación se realice de forma sesgada.

Existen tres teorías explicativas de la percepción del riesgo: el paradigma de medición axiomática, el sociocultural y el psicométrico. El primero de ellos consiste en la transformación de una información objetiva a una resolución subjetiva sobre la cual el individuo percibe el impacto que pueden tener determinados sucesos en su vida. El segundo paradigma estudia la forma en la que el entorno social condiciona al individuo en la valoración del riesgo. Por último, el *paradigma psicométrico*, que estudia el procedimiento cognitivo –utilización de heurísticos y factores emocionales– que sigue el individuo para la valoración del riesgo²¹. Este último explica el extremo temor de las personas hacia algunos peligros y su indiferencia hacia otros²².

Es preciso señalar que los jueces tienen el deber de conocer el derecho y saber aplicarlo, pero no tienen la obligación de ser expertos en la valoración del riesgo. Además, el período para evaluar la concurrencia de los tres peligros es escaso, y muchas veces se realiza en condiciones precarias, lo que da lugar a que esta valoración se desarrolle dependiendo del pensamiento intuitivo y experiencial. Por consiguiente, en este epígrafe estudiaremos la valoración del riesgo desde la perspectiva del *paradigma psicométrico*.

Según las teorías modernas de neurociencia, el procedimiento cognitivo que utiliza el ser humano para percibir, evaluar y controlar el riesgo está comprendido por dos sistemas diferentes. El primero de ellos es el denominado *sistema experiencial*, asociado a sentimientos que frecuentemente no tenemos presente. Se basa en la experiencia, acontecimientos que traen consigo una serie de emociones vividas –se caracteriza por su celeridad y por la utilización de heurísticos–. Por otro lado, nos encontramos con el calificado *sistema racional* o *analítico* mediante el cual se estima el riesgo a partir de parámetros lógicos, requiriendo un control de conciencia²³ –este proceso es más lento–.

²¹ STAJNOLOVIC, Milena. “Percepción social de riesgo: una mirada general y aplicación a la comunicación de salud”. *Revista de Comunicación y Salud*, vol. 5, 2015, p. 98.

²² SLOVIC, Paul, FISCHHOFF, Baruch y LICHTENSTEIN, Sarah. “The psychometric study of risk perception”, *Springer, Boston, MA*, vol.1, 1986, p. 4 y ss.

²³ SLOVIC, Paul et al. “Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality”. *Risk Analysis*, vol. 24, n. 2, 2004, p. 311.

Como ya hemos comentado anteriormente, el juez dispone de muy poco tiempo para valorar la concurrencia del *periculum*, lo que da lugar a que se utilice preferentemente el *sistema experiencial*. A continuación, estudiaremos este sistema y analizaremos la función que ejercen las emociones en la valoración del riesgo.

2.1. Sistema experiencial

Este sistema, como bien indica su nombre, se basa en la experiencia para la percepción del riesgo. La experiencia pasada es un factor importante para que el decisor perciba si existe o no el riesgo. Particularmente, “la experiencia del daño” es el factor más importante dado que es más fácil adoptar precauciones si previamente se ha materializado el daño²⁴. Por ejemplo, si en un accidente de tráfico salimos malheridos por no habernos puesto el cinturón, la próxima vez que subamos a un automóvil percibiremos la omisión de ponernos el cinturón como un riesgo elevado.

Esto es así, porque el riesgo que asociamos a un nuevo peligro lo comparamos con aquellos riesgos que nos son familiares²⁵. Por ello, la falta de experiencia podría entorpecer la valoración del riesgo por parte del decisor, provocando el *sesgo optimista*²⁶. Este sesgo aparece cuando el decisor considera que es menos vulnerable que otras personas porque la falta de experiencia le dificulta imaginar cómo podría afectarle²⁷. Siguiendo con este ejemplo, podría esperarse que un individuo que siempre ha llevado una conducción segura –sin ningún accidente– no utilice el cinturón de seguridad²⁸, traduciendo este riesgo en un “no me va a pasar”.

Podemos apreciar que la falta de experiencia puede obstaculizar la valoración del riesgo. Llevando esta última idea a la valoración del *periculum* por parte del juez, no es comparable la valoración de un juez que lleva ejerciendo cuarenta años a la de otro que apenas lleva unos pocos. El primero de ellos podría apreciar el riesgo de fuga, por ejemplo, basándose en casos anteriores similares, mientras que el segundo podría pasar este riesgo desapercibido por la falta de experiencia. Sabemos que la inteligencia artificial puede basarse en una infinidad de datos, pero ¿podría solventar la falta de experiencia del

²⁴ KELLER, Carmen, SIEGRIST, Michale y GUTSCHER, Heinz. “The role of the affect and availability heuristics in risk communication”. *Risk Analysis*, vol. 26, n. 3, 2006, p. 632.

²⁵ *Ibid.*

²⁶ WEINSTEIN, Neil D. “Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample”. *Journal of behavioral medicine*, vol. 10, n.5, 1987, p. 481 y ss.

²⁷ JOFFE, Hélène. “Risk: From perception to social representation”. *British journal of social psychology*, vol. 42, n.1, 2003, p. 57.

²⁸ SLOVIC, Paul, LICHTENSTEIN, Sarah, FIRSCHHOFF, Baruch. “Accident probabilities and seat belt usage: A psychological perspective”. *Accident Analysis & Prevention*, vol. 10, n. 4, 1978, p.283.

juez? Aunque la inteligencia artificial pueda tener en cuenta multitud de casos no sabremos qué es lo que ha tenido en cuenta para llegar a ese resultado –*black box*–, a diferencia de un juez con experiencia que éste sí podría motivar su decisión –aunque veremos más adelante que esta experiencia también plantea problemas–. Por consiguiente, una herramienta que emplea inteligencia artificial no puede sustituir a un juez con experiencia, pero sí podría auxiliar a uno que carezca de ella, pudiendo disminuir el *sesgo de optimismo*.

Una de las piezas centrales del *sistema experiencial* es la utilización de heurísticos, estrategias que utiliza el ser humano en el procesamiento de información para tomar una decisión. Estos implican una limitación en nuestro cerebro para procesar información debido a que nos vemos condicionados por las nuestras emociones y juicios de valor²⁹.

De forma distinta, una inteligencia artificial durante el proceso de información se basa en la totalidad de aspectos que plantea el problema, y no únicamente en aquellos que el decisor considera más relevantes³⁰. Como ya hemos adelantado, este sistema se caracteriza por la utilización de heurísticos, particularmente el heurístico de afección y el de accesibilidad o disponibilidad. A continuación, explicaremos ambos más detenidamente.

2.1.1. Heurístico de accesibilidad

Cuando valoramos un riesgo, estimamos la probabilidad de que un daño se materialice según la facilidad que tengamos en recordar determinados sucesos o asociaciones a los que podamos acceder con mayor facilidad³¹. Un ámbito muy claro donde podemos valorar la presencia de este heurístico es en como percibimos la frecuencia de diversas causas de muerte. Un ejemplo evidente que podemos apreciar en la actualidad es la muerte a causa la COVID-19. Por un lado, nos encontramos con personas que perciben un riesgo elevado de muerte respecto a esta enfermedad y, por otro lado, aquellas que les es indiferente porque no aprecian el riesgo de que esta enfermedad pueda causar la muerte. Esto es porque existen dos *proxies* en el heurístico de

²⁹ PÁEZ, Andrés. “Los sesgos cognitivos y la legitimidad racional de las decisiones judiciales”. *Filosofía y psicología, Universidad Externado de Colombia*, 2020, p.3.

³⁰ HIRSCHLAFF, Alan et al. “Inteligencia artificial y el rol del Juez: ¿el fin de los jueces convencionales o la confirmación de su insustituibilidad?”. *XLI Congreso Colombiano del Derecho Procesal*, 2020, p. 2.

³¹ NOVO, Mercedes, ARCE, Ramón, FARIÑA, Francisca. “El heurístico: perspectiva histórica, concepto y tipología”. *Jueces: Formación de juicios y sentencias*, 2003, p. 56.

accesibilidad: el daño directo y el daño indirecto³². El primero de ellos hace referencia a aquellos individuos que aprecian la enfermedad como un riesgo elevado de muerte porque algún amigo o familiar ha fallecido a causa de la COVID-19. El daño indirecto, por otro lado, influye en función de la frecuencia en la que el individuo haya oído hablar sobre esta enfermedad –medios de comunicación–.

El problema de este heurístico es que el daño directo podría provocar en el decisor la percepción de un riesgo mayor de lo que es realmente. Por ejemplo, supongamos que el juez en el caso X decidió no adoptar prisión preventiva como medida cautelar porque no apreció el riesgo de fuga y el reo finalmente acabó fugándose. La utilización de este heurístico va a dar lugar que el juez en el caso Y –posterior al caso X– adopte con mayor predisposición la prisión preventiva como medida cautelar porque acabará percibiendo el peligro de fuga como más fuerte de lo que es en realidad, dando lugar al fenómeno de *correlación ilusoria*.

La *correlación ilusoria* es entendida como un sesgo de accesibilidad que consiste en:

“Un error sistemático en el informe de tales relaciones. Es definida según el informe dado por un observador de una correlación entre dos clases de eventos que en realidad (i) no están correlacionados, o (ii) están correlacionados en menor medida que lo manifestado, o (iii) están correlacionados en el sentido contrario del que se manifiesta en el informe”³³.

En consecuencia, el juez estaría relacionando dos variables que o no tienen relación alguna, o bien, están correlacionados en una menor medida –riesgo de fuga del reo en caso X y riesgo de fuga del reo en caso Y–.

Como vemos, acudir a este heurístico para la valoración del riesgo presenta problemas. La inteligencia artificial, por otra parte, podría tener en consideración suficientes casos como para establecer una correlación objetiva entre ellos en cuestión de segundos, evitando que la evaluación del riesgo se realice únicamente con los casos a los que el juez tiene más fácil acceso en su mente y con los que quizás establecería relaciones erróneas.

³² PACHUR, Thorsten., HERTWIG, Ralph, STEINMANN, Florian. “How do people judge risks: availability heuristic, affect heuristic, or both?”. *Journal of Experimental Psychology: Applied*, vol.18, n.3, 2012, p. 315.

³³ CHAPMAN, Loren J. “Illusory correlation in observational report”. *Journal of Verbal Learning and Verbal Behavior*, vol. 6, n.1, 1967, cit. p. 194.

2.1.2. Heurístico de afecto

Este heurístico está estrechamente relacionado con el heurístico de disponibilidad. Las imágenes que nos vienen a la mente cuando recordamos un acontecimiento están etiquetadas con reacciones emocionales. Recordemos, por ejemplo, las imágenes de una furgoneta blanca que atravesaba la Rambla de Barcelona el 17 de agosto de 2017. Podemos apreciar cómo el recuerdo de estas imágenes evoca sentimientos de tristeza, dolor, miedo e incluso ira.

Acudimos a un “fondo afectivo” que contiene todas las etiquetas emocionales asociadas a ese recuerdo, y lo utilizamos –consciente o inconscientemente– para tomar una decisión³⁴. Concretamente, el sentimiento de temor es el que mayoritariamente determina la apreciación del riesgo por parte del decisor³⁵. Por ejemplo, supongamos que un juez debe apreciar el riesgo de reiteración delictiva de un reo negro que vive en un barrio peligroso. La apariencia del reo y el concepto “barrio peligroso” pueden evocar en el juez sentimientos negativos –temor– y por ello considerar que exista un riesgo mayor de reiteración delictiva del que verdaderamente hay.

El problema de este heurístico es que los componentes afectivos pueden ser fácilmente manipulables y, por ello, puede resultar una valoración del riesgo defectuosa³⁶. Asimismo, acudir a este heurístico para la valoración del *periculum* puede implicar una percepción del riesgo sesgada puesto que la decisión se basa en prejuicios propios del juez.

2.2. La influencia de las emociones en la valoración del riesgo

Este punto está especialmente relacionado al heurístico al que hemos aludido anteriormente, el heurístico de afección. Aunque, como ya hemos comentado, la utilización de este heurístico va aparejada a sesgos y prejuicios, las emociones que apreciamos en éste también son esenciales para la valoración del riesgo.

Para entender la importancia de las emociones, analizaremos brevemente la hipótesis del marcador somático³⁷. Esta hipótesis se basa en que, en el momento de tomar

³⁴ FINUCANE, Melissa et al. “The affect heuristic in judgments of risks and benefits”. *Journal of behavioral decision making*, vol. 13, n. 1, 2000, p. 3.

³⁵ SLOVIC, Paul et al. “Affect, risk, and decision making”. *Health psychology*, vol. 24, n. 4S, 2005, p. 3.

³⁶ SLOVIC, Paul et al. “Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality”. *Risk Analysis*, vol. 24, n. 2, 2004, p. 313.

³⁷ SLOVIC, Paul et al. “Affect, risk, and decision making”. *Health psychology*, vol. 24, n. 4S, 2005, p. 5.

una decisión, lo que hacemos es acudir a emociones ya vividas en situaciones concretas para poder dar respuesta a las situaciones actuales.

Para probar esta hipótesis, se estudió a aquellos individuos que tenían un daño cerebral y no podían asociar emociones a las posibles consecuencias que tendrían sus acciones³⁸. Tal y como hemos señalado anteriormente, las imágenes que tenemos en nuestra memoria están etiquetadas por sentimientos. Pues bien, estos sentimientos están vinculados a estados somáticos –positivos o negativos–, y cuando un somático negativo está vinculado a una imagen de un resultado futuro salta una alarma: la percepción del riesgo³⁹. Por ello, un individuo que no puede experimentar emociones no puede percibir el riesgo de forma correcta, no es capaz de dar una respuesta anticipada. Por ejemplo, imaginemos que durante nuestra formación nos encontramos con un profesor excelente y agradable en sus explicaciones. Cuando tengamos ocasión de volver a escoger este profesor, se activarán nuestros marcadores somáticos positivos reeligiéndolo antes que a cualquier otro. De lo contrario, si se tratase de un profesor nefasto, nuestros marcadores somáticos negativos harían saltar una alarma para no volver a escogarlo. Por tanto, podemos observar la importancia de las emociones para la apreciación del riesgo.

No obstante, cabe señalar que los marcadores somáticos nos hacen escoger una opción u otra a partir de lo que sentimos, pero esto no quiere decir que la decisión que tomemos sea la correcta. Supongamos, por ejemplo, un juez que está sesgado por experiencias desfavorables con individuos de raza negra y debe valorar el *periculum libertatis* de un reo negro. Se activarán los estados somáticos negativos haciendo que perciba un riesgo sobredimensionado o inexistente. Por ello, para realizar una correcta valoración deberá acudir a un estudio objetivo del riesgo –sistema analítico–.

En suma, podemos considerar que la valoración del riesgo está influenciada por las emociones. Una inteligencia artificial podría aprender la forma en la que se actuó en situaciones anteriores para dar respuesta a situaciones actuales, pero de ninguna forma podrá experimentar emociones. Sin embargo, sí que podría ser de ayuda cuando un juez se vea demasiado influenciado por sus emociones en el momento de la apreciación del

³⁸ DAMASIO, Antonio. *El error de Descartes. La emoción, la razón y el cerebro humano*. Barcelona, 2013, p. 256 y ss.

³⁹ FINUCANE, Melissa et al. “The affect heuristic in judgments of risks and benefits”. *Journal of behavioral decision making*, vol. 13, n. 1, 2000, p. 2.

periculum –aunque adelantamos, que un algoritmo también podría plantear una valoración sesgada–.

3. La prueba del *periculum libertatis* y la inteligencia artificial

El *periculum libertatis* es uno de los tres presupuestos para la adopción de medidas cautelares. El *periculum libertatis* se concreta en “*cualquier acción que pueda realizar el reo estando en libertad, y que pueda de algún modo comprometer la tutela que se dispense en el proceso*”⁴⁰. Está formado por tres posibles peligros: peligro de fuga, peligro de destrucción de pruebas y peligro de reiteración delictiva.

El objeto de este apartado es averiguar la posibilidad de que una herramienta que emplea inteligencia artificial pueda ayudarnos a obtener la probabilidad de que se genere un determinado comportamiento del reo en el futuro –fuga, obstaculización de la investigación mediante la destrucción de pruebas, o bien, reiteración en el hecho delictivo –. Trataremos también las herramientas existentes que actualmente están siendo utilizadas para la valoración del riesgo, concretamente, el riesgo de reiteración delictiva.

3.1. Evaluación del riesgo y la inteligencia artificial

Ante la evaluación del riesgo de fuga, destrucción de pruebas o reiteración delictiva nos encontramos con el problema de que éstos no pueden ser probados en cuanto a que no es posible constatar hechos que aún no se han producido⁴¹. Por ello, al tratarse de hechos prospectivos, lo que deberá probarse es una hipótesis probabilística.

Para poder probar esta hipótesis son necesarios los métodos estadísticos. Estos métodos se caracterizan por tener como objeto la creación de grupos prototípicos de criminales en los que se valoran ciertas características que, en función de su intensidad, determinan una mayor probabilidad de reiteración delictiva, fuga o destrucción de pruebas⁴². Estos métodos pueden ser trasladables a un mecanismo que utilice inteligencia artificial, pudiéndose sostener a partir de variables estadísticas introducidas por un programador. Veámoslo con respecto a cada *periculum*.

⁴⁰ NIEVA FENOLL, Jordi. *Derecho procesal III*. Marcial Pons, Madrid, 2017, p. 260.

⁴¹ MORA, Jeffry José. “Predictibilidad conductual y proceso penal: algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares”. *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, n. 2, 2021, p. 50.

⁴² ROMERO, Carlos María. “Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad”. *R.E.D.S.*, n.13, 2018, p. 42.

3.1.1. Riesgo de fuga

Antes de entrar a analizar las características para tener en cuenta en la valoración del riesgo de fuga, debemos señalar se trata de un peligro procesal, puesto que en el proceso penal se exige la presencia del reo para poder ser juzgado –o de lo contrario, se declararía su rebeldía con la consiguiente suspensión de los trámites procesales⁴³–.

Estudiando este riesgo desde la perspectiva de la adopción de la prisión provisional como medida cautelar, la LECrim concreta una serie de criterios objetivos para apreciar la concurrencia de este riesgo:

“Se atenderá conjuntamente a la naturaleza del hecho, a la gravedad de la pena que pudiera imponerse al imputado, a la situación familiar, laboral y económica de éste, así como a la inminencia de la celebración del juicio oral, en particular en aquellos supuestos en los que procede incoar el procedimiento para el enjuiciamiento rápido regulado en el título III del libro IV de esta ley.”⁴⁴

La hipótesis probabilística que probar en esta tipología de riesgo sería la pretensión de que el reo llegase a fugarse, por ejemplo: “José, puesto en libertad, se fugará y pondrá en peligro el objeto del proceso penal”. El juez deberá acreditar los medios por los que el reo podría evitar la acción de la justicia. Deberá razonar, en este caso, que “José no tiene domicilio habitual en España”, lo que implicaría la carencia de un vínculo suficientemente fuerte como para no huir del lugar de la celebración del juicio.

Asimismo, la doctrina creada por la jurisprudencia ha reafirmado los criterios establecidos por la LECrim y ha tenido en consideración otros: (i) la intensidad del gravamen de la pena; (ii) si ya ha habido una desobediencia previa hacia las autoridades judiciales –lo que implicaría que en un futuro no se respeten las nuevas resoluciones que vayan a ser dictadas⁴⁵–; (iii) las condiciones de arraigo –tales como el número de hijos, vecindad conocida, trabajo estable, reputación, etc.⁴⁶–; (iv) que el reo ya haya intentado fugarse⁴⁷ y (v) que el reo pertenezca a una organización criminal con ayuda de la cual pueda facilitar su huida⁴⁸. Por otro lado, el TC siguiendo la doctrina del TEDH, también

⁴³ A excepción de los supuestos contemplados en el artículo 786 de la LECrim.

⁴⁴ Artículo 503.1.1º a) LECrim

⁴⁵ Sentencia del Tribunal Supremo Sala de lo Penal, Sección 1ª, de 17 de mayo 2018/2257.

⁴⁶ Auto Juzgado de instrucción núm. 2 de Reus de 23 de marzo 2020/510.

⁴⁷ CAMACHO, Julia. “El fiscal pide que vuelva a prisión el guardia civil de “La manada” por riesgo de fuga”. *El periódico*, 28 de Junio de 2018. (disponible en <https://www.elperiodico.com/es/sociedad/20180628/el-guardia-civil-de-la-manada-quiso-sacarse-un-pasaporte-al-quedar-en-libertad-6913116>; última consulta 14/03/2021).

⁴⁸ Auto Juzgado Central de Instrucción de 30 de octubre 2014/1197.

ha hecho hincapié en la necesidad de tener en consideración el momento procesal para la apreciación del riesgo de fuga:

“Si en un primer momento cabría admitir que para preservar los fines constitucionalmente legítimos de la prisión provisional su adopción inicial se lleve a cabo atendiendo solamente al tipo de delito y a la gravedad de la pena, el transcurso del tiempo modifica estas circunstancias y por ello en la decisión de mantenimiento de la medida deben ponderarse inexcusablemente los datos personales del preso preventivo así como los del caso concreto.”⁴⁹

Podemos apreciar, entonces, que en un momento procesal posterior no bastará únicamente en reparar en el tipo de delito y la gravedad de la pena, sino que habrá que estudiar el individuo y el caso en cuestión. Asimismo, y teniendo en consideración el artículo anteriormente mencionado de la LECrim, podemos considerar que los criterios que se establecen implican una objetivación del riesgo. Es decir, se presume que la concurrencia de estos criterios, en mayor o menor medida, da lugar al riesgo de fuga.

Partiendo de esta idea, podríamos introducir los criterios que establece la LECrim y la doctrina jurisprudencial en un algoritmo como parámetros para la valoración del riesgo de fuga. Éste podría aprender a partir de casos anteriores la intensidad en la que se ha tenido en cuenta cada criterio, objetivando el riesgo de fuga –aunque, ya avanzamos que esto no implica necesariamente que el reo vaya a fugarse–.

3.1.2. Riesgo de destrucción de pruebas

El riesgo de destrucción, ocultación o alteración de pruebas es también un peligro procesal en cuanto a que de ser efectiva, frustraría uno de los fines del proceso penal, el correcto establecimiento de la verdad, puesto que implicaría alejar al juez de la efectividad de los hechos⁵⁰.

La hipótesis que probar para la valoración de este riesgo sería, por ejemplo, el supuesto de hecho “José, estando en libertad, hará desaparecer los rastros del delito cometido y pondrá en peligro el objeto del proceso penal privándolo de material probatorio”. En este supuesto, al igual que en el anterior, el juez deberá probar los medios de los que dispone el reo para proceder a dicha destrucción: “José, ostenta una posición de poder en una organización criminal”. En este supuesto de hecho, se tendría en consideración que José, al pertenecer a una organización criminal, puede alentar a

⁴⁹ Sentencia del Tribunal Constitucional del 15 de abril 62/1996.

⁵⁰NIEVA FENOLL, Jordi. *Derecho procesal III*. Marcial Pons, Madrid, 2017, p. 260.

terceros pertenecientes a dicha organización a la destrucción de pruebas o bien a la influencia sobre otros investigados, testigos o peritos. Una de las notas características de este riesgo es el factor del tiempo puesto que su mero transcurso puede implicar la desaparición total de este riesgo.

Para la valoración de este riesgo las características que se suelen tener en consideración son: (i) que el reo hubiese dificultado e impedido las medidas de investigación previamente⁵¹ –actuación que fácilmente puede reiterar–; (ii) la pertenencia a una organización criminal⁵²; (iii) el carácter reciente del hecho delictivo y pruebas que puedan desaparecer con facilidad.

Al igual que en la valoración del riesgo de fuga, también se podría utilizar la inteligencia artificial para valorar el riesgo de destrucción de pruebas. No obstante, como es el más reciente de los riesgos, los antecedentes serían escasos⁵³. La base de datos sería considerablemente menor a la del riesgo de fuga, por lo que no se podría establecer una intensidad adecuada de los criterios establecidos.

3.1.3. Riesgo de reiteración delictiva

El riesgo de reiteración delictiva es sin duda el riesgo a valorar más polémico. Se trata de una valoración gravemente contraria al derecho de presunción de inocencia puesto que presupone que el reo vuelva a cometer un delito⁵⁴. Este riesgo también se toma en consideración cuando se trata de otorgar beneficios penitenciarios o para establecer medidas de vigilancia policial una vez el individuo haya salido de prisión⁵⁵. Sin embargo, no es comparable la valoración del riesgo de reiteración de un condenado a la de un reo, que aún no se ha demostrado su imputabilidad y que, por tanto, es inocente⁵⁶.

La hipótesis para la valoración de este riesgo sería, entonces, “José, estando en libertad, volverá a delinquir”. En este caso, a diferencia del riesgo de fuga y destrucción de pruebas, el juez no parte factores objetivos –independientes de la consideración de la imputación– para la valoración de este riesgo, como el arraigo del reo en el país o el carácter reciente del hecho delictivo. Es todo lo contrario, partimos de la base de que un

⁵¹ Auto Juzgado Central de Instrucción del 16 de octubre 2017/252463.

⁵² Auto Juzgado Central de Instrucción nº 6 del 3 de marzo 0000096/2017.

⁵³ NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018, p. 65

⁵⁴ NIEVA FENOLL, Jordi. *Derecho procesal III*. Marcial Pons, Madrid, 2017, p. 293.

⁵⁵ NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018, p. 66.

⁵⁶ *Ibid.* p. 67.

individuo que presuntamente cometió un delito lo volverá a hacer, lo que constituye una presunción de culpabilidad⁵⁷.

Actualmente, están siendo utilizadas herramientas que emplean inteligencia artificial con el fin de “predecir” y evitar el delito. A continuación, estudiaremos algunas de éstas para revelar las consecuencias que puede implicar su utilización: *VioGén*, validada en España para la predicción de violencia contra las mujeres; el programa *COMPAS*, popular en Estados Unidos y el *Strategic Subject list*, utilizado por la policía de Chicago.

3.1.3.1. *VioGén*

El Sistema de Seguimiento Integral en los casos de Violencia de Género (Sistema VioGén), empezó su funcionamiento en España en 2007, en cumplimiento de lo establecido en la Ley Orgánica 1/2004, de 28 de diciembre de Medidas de Protección Integral contra la Violencia de Género⁵⁸.

El Sistema VioGén utiliza un algoritmo de valoración del riesgo que utiliza inteligencia artificial y Big Data. Esta herramienta se basa en el análisis permanente del riesgo, teniendo en consideración la peligrosidad del agresor, vulnerabilidad de la víctima y los factores que les rodean⁵⁹.

Este sistema parte de un primer cuestionario de valoración policial del riesgo contra la mujer –VPR–, teniendo en consideración 39 ítems de tipo histórico. Esta valoración deberá ser cumplimentada por la autoridad policial que dirija el caso en cuestión. A continuación, el sistema asigna uno de los cinco niveles de los que dispone: “no apreciado”, “bajo”, “medio”, “alto” o “extremo”, teniendo, el agente que ha cumplimentado el cuestionario, la potestad de modificar esta valoración al alza⁶⁰ –destacar que, la utilización de este sistema va dirigido exclusivamente a ayudar al agente en la apreciación del riesgo, en ningún momento lo que se pretende es la sustitución de este–. El resultado de esta valoración se incluye en el atestado policial que debe ser comunicado a la autoridad judicial correspondiente. Posteriormente, se realiza un

⁵⁷ NIEVA FENOLL, Jordi. *Derecho procesal III*. Marcial Pons, Madrid, 2017, p. 293 y ss.

⁵⁸ Sistema VioGén, Servicios al CIUDADANO. Recuperado de <http://www.interior.gob.es/web/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen>

⁵⁹ ESTÉVEZ, Lucana. “Inteligencia artificial y violencia contra las mujeres: ¿funcionan los sistemas automatizados de evaluación del riesgo?” *PERSPECTIVAS Revista de ciencias jurídicas y políticas*, n. 3, 2020, p. 132.

⁶⁰ *Ibid.*, p. 133.

segundo cuestionario de valoración policial para analizar la evolución del riesgo (VPER), teniendo en consideración esta vez 43 ítems.

Aunque esta herramienta disponga de grandes ventajas en cuanto a que hace una valoración teniendo en cuenta multitud de datos facilitando un análisis probabilístico y preventivo, trae también desventajas consigo. Debemos hacer especial referencia a una sentencia de la Audiencia Nacional dictada en 2020⁶¹.

Esta sentencia estimaba la responsabilidad patrimonial del Estado derivada del asesinato de Stefany González Escarramán a manos de su marido. En 2016, Stefany solicitó una orden de protección contra su marido como víctima de violencia de género. El mismo día es citada ante el Juzgado de Primera Instancia e Instrucción para la celebración de juicio rápido, dictando un auto de denegación de la orden de protección. La valoración inicial del riesgo del sistema VioGén fue de un nivel “no apreciado”, un resultado no acertado que tuvo como consecuencia el asesinato de la denunciante.

El sistema VioGén no apreció el riesgo que, aplicando una perspectiva de género y mediante una entrevista personal sí se hubiera apreciado. El asesinato de Stefany fue fruto de la falta de conciencia acerca de la violencia de género por parte de las autoridades y la confianza depositada en este algoritmo. Por ello, en este sentido, es de criticar la actuación policial. No tuvieron en cuenta otros factores que derivaban de las diligencias practicadas en el juicio que implicaban que Stefany se encontraba en un peligro mayor al que determinaba el sistema VioGén. A esto, se le sumó la incompetencia del juez de instrucción y la fiscalía por no haber impedido este resultado, puesto que se limitaron a seguir la valoración realizada por esta herramienta⁶².

Con todo esto podemos apreciar el riesgo que podría provocar la automatización de este tipo de decisiones. Es necesario en todo momento la intervención humana, el criterio de la autoridad policial y judicial para evitar que se cometan errores de tal magnitud.

⁶¹ Sentencia Audiencia Nacional, Sala de lo Contencioso Administrativo, Sección 5 de 20 de septiembre 2350/2020.

⁶² ESTÉVEZ, Lucana. “Inteligencia artificial y violencia contra las mujeres: ¿funcionan los sistemas automatizados de evaluación del riesgo?” *PERSPECTIVAS Revista de ciencias jurídicas y políticas*, n. 3, 2020, p. 136.

3.1.3.2. COMPAS

El software COMPAS⁶³ es un algoritmo desarrollado por Northointe Inc, popular en Estados Unidos y ampliamente utilizada para la evaluación del riesgo. Analiza el riesgo de reincidencia dentro de los dos años siguientes a su utilización a partir de 137 ítems relacionados con el reo y sus antecedentes penales⁶⁴.

Este algoritmo fue fomentado por el Estado de Wisconsin a través de la sentencia de la *Supreme Court* de Wisconsin, *State v. Loomis* en la que apreciaba la constitucionalidad del uso de algoritmos. En 2013, Eric Loomis fue condenado por saltarse un control policial mientras conducía un coche robado. Posteriormente, el tribunal de instancia consideró:

*“You’re identified, through the COMPAS assessment, as an individual who is at high risk to the community. In terms of weighing the various factors, I’m ruling out probation because of the seriousness of the crime and because your history, your history on supervision, and the risk assessment tools that have been utilized, suggest that you’re extremely high risk to re-offend”*⁶⁵.

El tribunal se sirvió de esta herramienta apreciando un riesgo alto de reiteración. La defensa de Eric Loomis decidió recurrir alegando la violación del derecho al debido proceso –*due process*–, sustentándose en tres argumentos. El primero de ellos era que al no conocer de forma precisa el funcionamiento del programa debido a la característica de propiedad privada –secreto comercial–, privó a la defensa la posibilidad de cuestionar la fiabilidad de esta herramienta. El segundo argumento se basaba en la violación del derecho a una sentencia individualizada en cuanto a que el programa valoraba el riesgo de reincidencia con datos basados en estadísticas de grupo. Por último, se argumentó que la utilización de este programa constituía una discriminación por razón de género debido a que era uno de los factores tomados en consideración por el sistema⁶⁶. Pese a estos argumentos, el tribunal ratificó la sentencia desestimando las pretensiones de la defensa, razonando que para la valoración del riesgo se tuvo en consideración otros factores y no únicamente la evaluación del sistema COMPAS. Pese a la aceptación de la utilización de este sistema, también impuso cautelas que deberán ser tenidas en la utilización de

⁶³ Perfiles de Gestión de Delincuentes Correccionales para Sanciones Alternativas (*Correctional Offender Management Profiling for Alternative Sanctions*).

⁶⁴ DRESSEL, Julia y FARID, Hany. “The accuracy, fairness, and limits of predicting recidivism”. *Science Advances*, vol. 4, n. 1, 2018, p. 1.

⁶⁵ Sentencia del 13 de julio de 2016: *State v. Loomis*, 881, Wisconsin.

⁶⁶ ROMERO, Carlos María. “Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad”. *R.E.D.S.*, n.13, 2018, p. 49.

sistemas similares de valoración del riesgo. Estableció que únicamente podrían utilizarse para:

- Imponer medidas alternativas de prisión a delincuentes que supongan un riesgo bajo de reiteración delictiva.
- Valorar si es posible la supervisión de forma efectiva y segura del acusado.
- Determinar la suspensión de la pena y de la libertad condicional, además de las condiciones de su incumplimiento⁶⁷.

Por último, este tribunal aclara que de ninguna forma este tipo de herramientas podrán ser utilizadas para determinar si el individuo evita o no la condena a prisión. A raíz de este caso, se produjo la admisibilidad de los algoritmos de valoración del riesgo de reincidencia en Estados Unidos.

Además de las vulneraciones que, como hemos visto, conlleva la utilización de esta herramienta para la evaluación del riesgo, un estudio reciente de ProPublica ha detectado que también es racista, en tanto que establece índices de peligrosidad más altos en sujetos de raza negra que en personas de raza blanca⁶⁸.

Teniendo en consideración el período de tiempo en el que predice el riesgo de reincidencia –2 años posteriores a su utilización–, este estudio descubrió:

- Que el programa asociaba un mayor peligro de reincidencia a los individuos negros de lo que realmente tenían. Los individuos afroamericanos que no reincidieron durante el período de dos años tuvieron casi el doble de probabilidades de ser calificados como de mayor riesgo en comparación con los individuos de raza blanca (45% contra un 23%).
- Que se evaluaban a los individuos de raza blanca con un menor índice de peligro del que verdaderamente representaban. Los individuos de raza blanca que reincidieron durante los dos años siguientes fueron calificados con un riesgo bajo, teniendo casi el doble de reincidencia respecto a los sujetos negros (48% contra un 28%).

⁶⁷ BORGES, Raquel. “El sesgo de la máquina en la toma de decisiones en el proceso penal”. *Ius et scientia*, vol. 6, n. 2, 2020, p. 64.

⁶⁸ NIEVA FENOLL, Jordi. *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018, p. 69.

- Que incluso teniendo el programa en consideración criterios como la edad, antecedentes penales, y género, era más probable que los acusados negros tuviesen un resultado de riesgo alto frente al resto (45% más).
- Que los individuos negros tenían el doble de probabilidad de ser evaluados con un mayor riesgo de reiteración violenta que los de raza blanca, que tenían una probabilidad del 63% de ser clasificados de riesgo bajo.
- Que incluso teniendo en consideración los criterios de edad, antecedentes penales, y género, los acusados negros tenían un 77% más de posibilidades de ser clasificados con un riesgo más alto de reincidencia violenta que los de raza blanca⁶⁹.

A todo ello, el departamento de investigación de Northpointe Inc., cuestionó el análisis de los puntos anteriormente explicados⁷⁰. El estudio se centró en la justificación del hecho de que la evaluación de riesgo refleja una realidad puesto que existe un porcentaje elevado de individuos de raza negra que han reincidido nuevamente. Por ello, justifica que no es sorprendente que estos individuos obtengan una valoración del riesgo más alta.

3.1.3.3. *Strategic Subject List –LSS–*

De nuevo, nos encontramos con una herramienta que con ayuda de la inteligencia artificial pretende anticipar la conducta delictiva. El *Strategic Subject List –Lista Estratégica de Sujetos–*, es utilizado por la policía de Chicago para determinar la probabilidad en la que un individuo pueda verse involucrado en un tiroteo u homicidio, ya sea como víctima o delincuente⁷¹.

Este algoritmo tiene consideración 8 ítems para la valoración del riesgo: (i) el número de veces que se ha sido víctima de un tiroteo; (ii) la edad del sujeto en el último arresto, (iii) el número de veces como víctima de alguna agresión; (iv) el número de arrestos por delitos violentos; (v) la pertenencia a una banda callejera; (vi) el número de arrestos por narcóticos; (vii) tendencia a la actividad delictiva y (viii) el número de veces

⁶⁹ LARSON, Jeff et al. “How We Analyzed the COMPAS Recidivism Algorithm”. *ProPublica*, 2016, p. 2.

⁷⁰ DIETERICH, William, MENDOZA, Christina y BRENNAN, Tim. “COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity”. *Northpointe Inc. Research Department*, 2016, p. 4 y ss.

⁷¹ ASHER, Jeff y ARTHUR, Rob. “Inside the Algorithm That Tries to Predict Gun Violence in Chicago”. *The New York Times*, 13 de junio 2017 (disponible en <https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html>; última consulta 04/04/2021).

arrestado por tenencia ilegal de armas⁷². A partir de estos, el algoritmo produce una lista de sujetos a los que se le asigna un resultado numérico comprendido del 0 –valor mínimo– al 500 –valor máximo– que indica el riesgo en el que individuo se va a ver involucrado en un tiroteo u homicidio.

La intensidad en la que el algoritmo tiene en consideración estos ítems no es pública, pero un estudio del *New York Times*, concluyó que los factores que más ponderan son la edad del sujeto en el último arresto y el número de veces como víctima de alguna agresión⁷³.

El Departamento de Policía de Chicago utiliza esta herramienta con dos fines: intervención y aplicación. El primer propósito consiste en una notificación personalizada a aquellos sujetos que hayan obtenido una puntuación elevada. Por un lado, advertir a los individuos que tienen un mayor riesgo de delinquir que se encuentran bajo sospecha y que se enfrentan a sanciones mayores en caso de participar en actividades delictivas. Y, por otro lado, alertar a los individuos que estén en peligro⁷⁴. El segundo de los propósitos de la utilización de esta herramienta es su aplicación como recurso de investigación con sus correspondientes límites: podrá ser utilizada en el inicio de una investigación, pero un individuo no podrá ser arrestado por el simple hecho de aparecer en la lista que ofrece el SSL puesto que no es una herramienta para establecer una sospecha razonable⁷⁵.

Pese a las ventajas que podría brindarnos esta herramienta, podemos encontrar serios inconvenientes en su aplicación. En primer lugar, no se valora al individuo por sus acciones –que es lo que justificaría iniciar una investigación–, sino que se valora a partir de una base de datos que contiene un amplio historial de antecedentes penales⁷⁶. Asimismo, el hecho de que un individuo esté en esta lista no implica necesariamente que sea un delincuente potencial –podría ser una víctima potencial–, además, los factores de riesgo podrían cambiar y no verse reflejados en la lista⁷⁷. Por otro lado, desde el punto de vista de la posible víctima, la notificación de la policía de que se encuentra en peligro

⁷² TUCEK, Aron. “Constraining Big Brother: The Legal Deficiencies Surrounding Chicago's Use of the Strategic Subject List”. *U. Chi. Legal F*, v.2018, 2019, p. 427.

⁷³ *Ibid.*

⁷⁴ SHEEHEY, Bonnie. “Algorithmic paranoia: the temporal governmentality of predictive policing”. *Ethics and Information Technology*, vol21, n. 1, 2019, p. 53 y ss.

⁷⁵ TUCEK, Aron. “Constraining Big Brother: The Legal Deficiencies Surrounding Chicago's Use of the Strategic Subject List”. *U. Chi. Legal F*, v.2018, 2019, p. 427.

⁷⁶ FERGUSON, Andrew G. “Illuminating black data policing”. *Ohio St. J. Crim.*, vol. 15, n. 503, 2017, p. 520.

⁷⁷ FERGUSON, Andrew G. “Policing predictive policing”. *Wash. UL Rev.*, vol. 94, n. 5, 2017, p. 1139.

puede dar lugar a una alteración drástica en su vida, pudiendo llegar a casos de ansiedad o miedo a salir a la calle⁷⁸.

En cuanto a todo lo abordado con anterioridad, debemos apreciar el peligro que comporta la utilización de este tipo de programas para la valoración del riesgo, tanto desde la perspectiva de la víctima como la del acusado. La aplicación de estas herramientas podría implicar la provocación de un sesgo de anclaje, condicionando al decisor en su valoración. Lo vimos en 2016 cuando las autoridades policiales no tuvieron en consideración otros factores más que el resultado que proporcionó VioGén para la valoración del peligro que corría Stefany. Asimismo, a la hora de adoptar medidas cautelares, el juez podría hallarse influenciado por el resultado, omitiendo una valoración más profunda porque su intuición ha sido corroborada por un algoritmo. Podemos apreciar, además, los posibles sesgos que pueden suponer la utilización de estos algoritmos, como vimos con COMPAS, que resultó ser racista.

Por otro lado, a estas desventajas debemos sumarle el riesgo que puede comportar el mal uso de este tipo de herramientas por parte de los humanos. Es el caso del “sheriff” de Pasco –Florida–. Este individuo utilizaba un algoritmo, que al igual que el *Strategic Subject List*, determina una lista de sujetos que pueden volver a delinquir en un futuro. Es decir, persigue los delitos que aún no han sido cometidos –y que quizás nunca se cometan–. A partir de esta lista, el “sheriff” detenía a estos sujetos y acosaba a sus familiares antes de que “volviesen a delinquir”⁷⁹.

3.2. El problema del criterio de predicción

El problema que trae consigo la utilización de programas que emplean inteligencia artificial es el criterio de predicción. Realmente, no están prediciendo la probabilidad en la que una determinada acción acontezca –que el reo se fugue, destruya pruebas o reitere en el hecho delictivo–, sino que se evalúa el riesgo que existe de que se produzca el mismo⁸⁰. En otras palabras, no se trata de evaluar si el individuo posee la cualidad de peligroso, sino que más bien se trata de estudiar dentro de un contexto el conjunto de

⁷⁸ HERNÁNDEZ, María. “Inteligencia artificial y Derecho Penal”, *Actualidad Jurídica Iberoamericana*, 10 bis, 2019, p. 819.

⁷⁹ PROTO, Lucas. “El ‘Minority report’ de Florida: un ‘sheriff’ persigue delitos todavía no cometidos”. *El Confidencial*, 15 de marzo de 2021 (disponible en https://www.elconfidencial.com/mundo/2021-03-15/el-minority-report-de-florida-un-sheriff-persigue-delitos-todavia-no-cometidos_2993108/; última consulta 05/03/2021).

⁸⁰ JULIÀ, Miquel. “Proceso penal y (neuro)ciencia: una interacción desorientada” p. 129. Disponible en <http://diposit.ub.edu/dspace/handle/2445/143437>

factores – personales, ambientales...– que pueden abocar a la comisión de un delito, la fuga o la destrucción de pruebas, permitiendo efectuar un pronóstico sobre la conducta futura. Por tanto, en este apartado estudiaremos la predicción del riesgo y no la peligrosidad del individuo.

La psicología cognitiva actualmente utiliza principalmente dos modelos diferentes para la predicción del riesgo: el modelo clínico y el modelo actuarial. El primero de ellos consiste en una evaluación realizada según el criterio del profesional, que predecirá el riesgo en función de su formación, preferencias o hábitos profesionales⁸¹. El problema de esta metodología es que hay un elevado margen de subjetividad en cuanto a que el resultado dependerá de la competencia del profesional y de los posibles sesgos que éste tenga presentes.

Por otro lado, tenemos los modelos actuariales o estadísticos, que basan el pronóstico del riesgo en el estudio de un grupo prototípico que comparten una serie de características, eliminando el elemento subjetivo del anterior modelo. Se utilizan modelos estadísticos y matemáticos para evaluar, en este caso, el riesgo de reiteración delictiva, fuga o destrucción de pruebas. Las herramientas de predicción del riesgo que utilizan inteligencia artificial podrían ubicarse dentro de este modelo.

El problema que comporta la utilización de este modelo –y que ya hemos podido apreciar en la utilización de este tipo de herramientas– es que la valoración del riesgo recae en aspectos puramente objetivos, totalmente despersonalizados del individuo. No se llega a la examinación del riesgo del individuo en cuestión, sino que tiene por objeto su clasificación e integración en un grupo al que se le ha señalado un determinado grado de riesgo⁸².

Estos métodos no nos proporcionan una información causal, sino correlacional. Es decir, el hecho de que estén presentes determinados factores no implica que el reo actúe de una forma u otra. Por ejemplo, en el riesgo de fuga, el hecho de que el reo disponga de medios económicos no quiere decir que vaya a fugarse. Estas herramientas, al tratar de predecir el delito, la fuga o la destrucción de pruebas, lo que hacen es objetivar el riesgo, no avanzarse a la conducta.

⁸¹ PUEYO, Antonio y REDONDO, Santiago. “Predicción de la violencia: entre la peligrosidad y la valoración del riesgo de violencia”. *Papeles del Psicólogo*, vol. 28, n.3, p. 166.

⁸² *Ibid.*

4. Problemática del uso de la inteligencia artificial en la valoración del riesgo

De la misma forma en la que la utilización de la inteligencia artificial comporta ventajas considerables al ser humano, su empleo en el proceso penal también puede implicar problemas éticos y vulneraciones en los derechos fundamentales.

Tal y como ya hemos hecho referencia, cuando un algoritmo “aprende”, lo hace a partir de datos previamente introducidos por el ser humano. En este sentido, podría ser que los datos introducidos por el programador hayan sido recogidos de forma sesgada, teniendo como consecuencia la reproducción de estos fallos por el algoritmo. Dicho de otro modo, estos algoritmos serán racistas, sexistas y clasistas de la misma forma en la que lo son sus programadores⁸³.

A lo largo de este epígrafe estudiaremos los posibles desafíos que puede suponer la utilización de estas herramientas.

4.1. El sesgo de los datos

Encontramos dos posicionamientos diferentes ante la utilización de este tipo de programas. En primer lugar, están los que defienden su implementación con el fin de acabar con los sesgos subjetivos y facilitar el arduo trabajo de la valoración y predicción del riesgo. Por otro lado, están aquellos que prefieren rescindir de su utilización bajo el temor de no poder controlar el funcionamiento de estos programas⁸⁴. A continuación, estudiaremos uno de los problemas que plantea el uso de estos programas, siguiendo el criterio del segundo grupo de críticos.

Hallamos dos tipologías principales de sesgos en la utilización de algoritmos de inteligencia artificial: los sesgos en los datos de entrenamiento –*biased training data*– y, por otro lado, los sesgos por una distribución no uniforme en los datos de entrenamiento –*biased unequal ground truth*–⁸⁵.

El primero de ellos se encuentra en los algoritmos de aprendizaje automático. Recordemos que este algoritmo –a diferencia del algoritmo clásico– se entrena analizando

⁸³ BORGES, Raquel. “El sesgo de la máquina en la toma de decisiones en el proceso penal”. *Ius et scientia*, vol. 6, n. 2, 2020, p. 54.

⁸⁴ MIRÓ, Fernando. “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots”. *Revista de derecho penal y criminología*, vol. 3, n. 20, 2018, p.110.

⁸⁵ HECKER, Philipp. “Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law”. *Common Market Law review*, 2018, p. 5.

ejemplos y estableciendo patrones estadísticos. Pues bien, con la introducción de estos ejemplos y datos para su entrenamiento pueden aparecer dos posibles sesgos:

- Que el algoritmo sea entrenado con datos erróneamente calificados por el programador –que una imagen de un perro estuviese etiquetada como si fuese la imagen de un gato–.
- Que el conjunto de datos esté acotado, reduciendo la validez e incrementando el error de la interpretación de los resultados –que únicamente tuviese en consideración una determinada raza de perro, estaría sesgado a diferenciar únicamente esa raza–.

La presencia de estos sesgos en los algoritmos implica la discriminación de grupos marginados, implicando una vulneración en sus derechos⁸⁶. Por ello, conviene tener presente las limitaciones de un algoritmo de aprendizaje automático que derivan de estos sesgos. Para ello, sería necesaria la publicidad de estos –que no vimos en el caso *State v. Loomis*– con el fin de poder ejercer correctamente el derecho de defensa.

La segunda tipología de sesgos en la utilización de estos algoritmos, son a causa de la existencia de variables que en la realidad se distribuyen de manera desigual. En otras palabras, aquellos factores inherentes al individuo que el algoritmo tiene en consideración y que no tienen relación con el actuar del mismo, sino con el actuar de sus iguales (mismo género, edad, etnia...)⁸⁷. No obstante, ¿deberían retirarse estos factores del criterio de valoración del algoritmo? Es difícil no tener en consideración el factor de género en el momento de evaluar el riesgo de reiteración cuando hablamos, por ejemplo, de delitos contra la libertad e indemnidad sexuales, en cuanto a que la tasa delictiva es notoriamente diferente entre hombres y mujeres.

Teniendo en consideración lo anteriormente explicado, la utilización de estos algoritmos podría implicar la vulneración del derecho de no discriminación y defensa del individuo evaluado. Como hemos visto, en caso de que los datos introducidos presentasen algún tipo de sesgo, se le atribuiría al reo una responsabilidad penal derivada de la evaluación de la actuación de otras personas, sobre las que él no es responsable. Asimismo, al alimentar al algoritmo de datos sesgados que reflejan nuestros propios

⁸⁶ MIRÓ, Fernando. “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots”. *Revista de derecho penal y criminología*, vol. 3, n. 20, 2018, p.123.

⁸⁷ *Ibid*, p. 124

prejuicios, lo que realmente estaríamos haciendo sería perpetuar y arraigar el sesgo humano⁸⁸.

4.2. Uso ético de la Inteligencia Artificial

Uno de los riesgos éticos que surge a raíz de la utilización de las herramientas de inteligencia artificial es la falta de transparencia. Como comentamos al inicio de este trabajo, la inteligencia artificial opera como una *black box*. Sabemos los datos que entran, y los datos que salen –predicciones–, pero es incomprensible para la mente humana el proceso interno por el cual llega a estos datos de salida.

Esta problemática podría dar lugar a dos situaciones. La primera de ellas es que, el juez evaluador del riesgo, al no entender el proceso por el cual el algoritmo ha llegado a esa decisión, opte por no tener en cuenta esta recomendación dando un mayor peso a su experiencia personal. Esta situación daría lugar al malgasto de recursos en el desarrollo de herramientas que finalmente no van a ser utilizadas. En la segunda situación, el juez evaluador del riesgo podría considerar los resultados que le proporciona el algoritmo como suficientes para la evaluación del riesgo, omitiendo una interpretación propia de los resultados aportados por el programa. En este caso, el problema sería reducir el proceso de valoración a lo que dice el algoritmo, convirtiendo la valoración del riesgo en una decisión automatizada⁸⁹.

Otro de los problemas que plantea la falta de transparencia en la utilización de estos algoritmos en la valoración del riesgo, es que difícilmente la defensa podrá demostrar si los factores como el género o la raza pudieron influir en la puntuación determinada por el algoritmo.

Adicionalmente, cabe destacar que, aunque estas herramientas funcionen como una *black box*, aquellas empresas como Northpointe Inc. que las patentan, tienen un especial interés en envolver sus algoritmos en secreto para seguir siendo competitivas, apuntando hacia una forma más de negocio⁹⁰.

⁸⁸ BURANYI, Stephen. “Rise of the racist robots – how AI is learning all our worst impulses”. *The Guardian*, 8 de agosto de 2017 (disponible en <https://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses>; última consulta 16/03/2021).

⁸⁹ GONZÁLEZ-ÁLVAREZ, José Luis; Jorge SANTOS y Miguel CAMACHO, “Policía predictiva en España. Aplicación y retos futuros”. *Fundación Universitaria Behavior & Law*, vol. 6, n.1, 2020, p. 28.

⁹⁰ KEHL, Danielle, GUO, Priscilla y KESSLER, Samuel. “Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing”. *Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School*, 2017, p. 28.

4.3. Derechos del reo

Como ya vimos, en 2016 el Tribunal Supremo del Estado de Wisconsin dictó una sentencia en la que se adoptaba la admisión del uso de herramientas que utilizan inteligencia artificial en el proceso penal –concretamente, en la valoración del peligro de reiteración delictiva–. Al haber sido avalada la utilización de este tipo de herramientas, podemos decir que estamos ante una nueva era en la aplicación de la justicia. Con razón o sin ella, estas nuevas herramientas han llegado a los tribunales para quedarse.

A lo largo de este trabajo hemos podido ir apreciando que el uso de estas nuevas tecnologías que emplean inteligencia artificial para la valoración del riesgo incide en los derechos fundamentales del reo, aunque también vimos que podía afectar a la víctima – como bien pudimos observar en el caso de Stefany con la aplicación del programa VioGén–. En este punto, estudiaremos el impacto que puede tener la utilización de este tipo de herramientas desde la perspectiva de los derechos fundamentales del reo, derechos a los que no debe renunciar.

Concretamente, analizaremos cómo afecta este uso en los derechos que forman parte del contenido del debido proceso –o proceso con todas las garantías–. El debido proceso lo encontramos regulado en el artículo 24 CE:

“1. Todas las personas tienen derecho a obtener la tutela efectiva de los jueces y tribunales en el ejercicio de sus derechos e intereses legítimos, sin que, en ningún caso, pueda producirse indefensión.

2. Asimismo, todos tienen derecho al Juez ordinario predeterminado por la ley, a la defensa y a la asistencia de letrado, a ser informados de la acusación formulada contra ellos, a un proceso público sin dilaciones indebidas y con todas las garantías, a utilizar los medios de prueba pertinentes para su defensa, a no declarar contra sí mismos, a no confesarse culpables y a la presunción de inocencia.

La ley regulará los casos en que, por razón de parentesco o de secreto profesional, no se estará obligado a declarar sobre hechos presuntamente delictivos.”

Dentro de este artículo debemos hacer una breve distinción. Tanto en el primer como en el segundo apartado se consagra el derecho de la tutela judicial efectiva. El primer apartado asegura la tutela por medio del acceso al proceso, pero este derecho no se agota únicamente con acceder a los tribunales, sino que, como podemos observar en el segundo apartado, también se asegura la tutela efectiva mediante las llamadas “garantías

procesales”⁹¹. A continuación, veremos cómo la utilización de herramientas de inteligencia artificial para la valoración del *periculum* vulnera algunas de las garantías procesales establecidas en el artículo 24.2 CE.

4.3.1. Derecho a un juez imparcial

Esta garantía procesal no está contemplada de forma expresa en el artículo 24.2 CE, sino que fue a partir de la STC 145/1988⁹² que el TC determinó que este derecho debía incluirse dentro del propio artículo 24, considerando el derecho a la imparcialidad judicial una exigencia del debido proceso⁹³. Asimismo, encontramos la presencia de este derecho en diferentes tratados internacionales ratificados por España, como el artículo 6.1 CEDH, el artículo 10 DUDH y el artículo 14 PIDCP⁹⁴.

Este derecho consiste en el deber del juez, por un lado, de no posicionarse a favor o en contra de cualquiera de las partes del proceso – imparcialidad subjetiva– y, por otro lado, de renunciar a cualquier prejuicio que pueda hacer peligrar una valoración neutra y justa⁹⁵.

Para analizar el deber de imparcialidad objetiva del juez debemos diferenciar dos aspectos, el *contenido* de las decisiones y los *motivos* para decidir. El primero de ellos alude a la correcta aplicación de la ley –legalidad–, mientras que el segundo reside en las circunstancias que el juez ha tenido en cuenta para tomar la decisión –imparcialidad–⁹⁶. Debemos señalar que, aunque la decisión adoptada contenga un contenido correcto, si ésta ha sido tomada por motivos incorrectos, implicará la deslegitimación de la decisión. Por ejemplo, el caso en el que un juez, en el momento de hacer la valoración del *periculum libertatis* de un reo negro, valore un riesgo de fuga elevado justificando que el reo carece de domicilio habitual en España –artículo 503.1.1º a LECrim –, pero el motivo real que le ha llevado a apreciar este riesgo ha sido la presencia de un prejuicio racial. Por tanto, para poder garantizar este derecho, el juez debe hacer una valoración del riesgo libre de sesgos y parcialidades.

⁹¹ BERIAN, Iñigo de Miguel y PÉREZ, Miren Josune. “La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados”. *Revista de Derecho UNED*, n. 25, 2019, p. 547.

⁹² Sentencia Tribunal Constitucional del 12 de julio 145/1988.

⁹³ COLMENERO, Miguel. “La garantía del derecho a un juez imparcial”. *Persona y derecho*, vol. 55, 2006, p. 723.

⁹⁴ Sentencia del Tribunal Constitucional del 22 de marzo 39/2004.

⁹⁵ Sentencia del Tribunal Constitucional del 12 de julio 145/1988.

⁹⁶ AGUILÓ, Josep. “Imparcialidad y concepciones del derecho”. *Juridicas*, vol. 6, n. 2, 2009, p. 32.

La utilización de algoritmos podría conllevar a la vulneración de este derecho. El juez podría verse influenciado en los *motivos* para decidir puesto que podría basarse y confiar en el resultado ofrecido por el programa⁹⁷. Es decir, un juez tendrá mayor predisposición en adoptar prisión provisional si sus sospechas son apoyadas por un programa que valora el riesgo a partir de una gran multitud de datos y de forma objetiva. Estaremos, por tanto, ante un sesgo de anclaje, que consistirá en “anclar” al decisor en la primera información ofrecida.

Todo ello suponiendo que el algoritmo sea fiable y parta de una base de datos completamente objetiva. Como hemos visto anteriormente, podría darse el caso que el algoritmo utilizado para la valoración del riesgo sea entrenado con datos sesgados. En este caso, la imparcialidad del juez se verá afectada por los sesgos que traen consigo estos datos. Es decir, si toma en consideración el resultado sesgado, indirectamente, estará adhiriendo a sus *motivos* para decidir los sesgos que ha aprendido el algoritmo –tales como el género, etnia, nivel económico...–.

Por otro lado, debemos hacer especial referencia a la implicación de este derecho en la valoración del riesgo de reiteración delictiva. En el momento en el que el juez debe valorar este riesgo para la adopción de medidas cautelares, está partiendo de la idea de que el reo reincidirá en un futuro, posicionándose en contra de él y vulnerando la imparcialidad subjetiva ya sea con o sin algoritmos. No obstante, dentro del sistema penal español se intenta lidiar con esta parcialidad al adscribirse al sistema acusatorio mixto. Por ello, el juez de instrucción que, por ejemplo, ha apreciado un riesgo alto de reiteración delictiva no podrá ser parte del tribunal en la fase del juicio oral –principio acusatorio–, al verse posicionado parcialmente contra el reo⁹⁸. La dificultad se nos plantea cuando esta valoración se realiza para dictar sentencia, por ejemplo, para determinar la condena. Lo vimos en el caso *State vs Loomis*. El procedimiento penal de EEUU es diferente al español, consta de dos fases. En la primera de ellas se determina la culpabilidad y en la segunda se especifica la pena. En este caso, cuando se dispuso de la utilización del algoritmo para valorar el riesgo de reincidencia del reo, se estaba asumiendo que éste –culpable– volvería a delinquir en un futuro. El tribunal entonces, ya se estaba posicionando en contra del él al intentar predecir el futuro, haciendo depender la pena

⁹⁷ BERIAN, Iñigo de Miguel y PÉREZ, Miren Josune. “La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados”. *Revista de Derecho UNED*, n. 25, 2019, p. 531.

⁹⁸ Artículo 219.11ª LOPJ.

que se le impondría a partir de una premisa inexistente, unos hechos que aún no habían sucedido –y que quizás no sucederían–.

4.3.2. Derecho de defensa

El derecho de defensa lo encontramos regulado en el artículo 24 CE. Podemos analizar este derecho desde dos puntos de vista distintos. Por un lado, el primer apartado del artículo 24 CE prohíbe la indefensión, ligado al deber del juez de impedir alteraciones que puedan dar lugar a esta prohibición⁹⁹. En éste, encontramos consagrado el derecho de la acusación a llevar a los tribunales el asunto por el cual se han visto vulnerados sus derechos. Por otro lado, en el apartado segundo se establece el derecho de defensa, consistente en la posibilidad de que el reo disponga de los medios necesarios para poder hacer frente al inicio de un procedimiento contra él. El propósito de este derecho es garantizar los principios de igualdad entre partes y contradicción¹⁰⁰. Es decir, la posibilidad de que ambas partes conozcan los materiales de hecho y derecho que llevan al juez a tomar una decisión y, por otro lado, la posibilidad de poder rebatir esta decisión¹⁰¹.

Para que haya una vulneración del derecho de defensa, el Tribunal Constitucional ha dispuesto que tal vulneración debe de ir acompañada de un “*perjuicio real y efectivo en las posibilidades de defensa de quien las denuncie*”¹⁰². Del mismo modo, la indefensión material no es sólo respecto al artículo 24.1 CE, sino que también lo es respecto a los derechos reconocidos en el artículo 24.2 CE: “*derechos al juez ordinario predeterminado por la Ley y a la imparcialidad judicial, en relación con las incidencias en las composiciones de los órganos judiciales*”¹⁰³ así como el derecho al debido proceso¹⁰⁴.

Teniendo esto en consideración, la posibilidad de que el reo alegue indefensión por la utilización de algoritmos para la valoración del riesgo es considerablemente amplia. En primer lugar, por la falta de transparencia que conlleva el uso de algoritmos. Como hemos reiterado en varias ocasiones, no es posible conocer todos los aspectos en los que

⁹⁹ Sentencia del Tribunal Constitucional del 12 de junio 152/2000.

¹⁰⁰ CATENA, Víctor Manuel. “Sobre el derecho de defensa: cuestiones generales”. *Teoría y Derecho*, n. 8, 2010 p. 17

¹⁰¹ BERIAN, Iñigo de Miguel y PÉREZ, Miren Josune. “La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados”. *Revista de Derecho UNED*, n. 25, 2019, p. 552.

¹⁰² Sentencia del Tribunal Constitucional del 26 de septiembre 233/2005.

¹⁰³ Sentencia del Tribunal Constitucional del 12 de septiembre 215/2005.

¹⁰⁴ Sentencia del Tribunal Constitucional del 3 de julio 219/2006.

el algoritmo se ha apoyado para llegar a un resultado, lo que implica la imposibilidad material de que el reo pueda rebatir y aportar pruebas a su favor. A esto, debemos añadirle el secreto comercial que protege la estructura que forman estos algoritmos, lo que dificulta aún más su transparencia. Por ello, difícilmente el reo podrá rebatir el resultado que ha determinado el algoritmo. Por otro lado, y tal como hemos podido ver en la jurisprudencia del TC, la vulneración de cualquier garantía consagrada por el artículo 24.2 CE supone también la transgresión del derecho de defensa. Por consiguiente –y tal y como hemos podido apreciar en el apartado anterior–, el uso de algoritmos no solo quebranta el derecho de defensa por la falta de transparencia que comporta, sino que indirectamente, también lo hace al quebrantar otros derechos establecidos por el artículo 24.2 CE.

4.3.3. Derecho a la presunción de inocencia

El derecho a la presunción de inocencia lo vemos configurado en el artículo 24.2 CE como derecho fundamental. Se trata de un principio configurador del sistema penal que debe orientar al juez durante todo el proceso¹⁰⁵. La presunción de inocencia en sentido estricto determina que, dentro de un juicio penal, el reo es inocente hasta que se pruebe lo contrario, lo que implica que, en caso de duda, el juez deberá absolver, el *in dubio pro reo*. Esta última afirmación, sin embargo, no es del todo cierta puesto que difícilmente el juez puede estar completamente libre de dudas¹⁰⁶.

Antes de entrar a valorar la forma en la que este derecho puede verse afectado con el uso de los algoritmos, lo estudiaremos desde la perspectiva de la valoración del riesgo para la adopción de medidas cautelares. *A priori*, podríamos pensar que en la instrucción se vulnera el derecho a la presunción de inocencia. Y es cierto puesto que, de no haber una mínima duda de que el reo es culpable, no se podría iniciar una investigación. Estamos ante una vulneración inevitable y, por consiguiente, legítima. Por ello, en la valoración del *periculum libertatis* para la adopción de medidas cautelares vemos sacrificado este derecho, pero con el fin de salvaguardar otros bienes jurídicos igual de preciados. Podemos entender en cierto modo la vulneración de este derecho en la valoración del *periculum*, pero para evitar prolongar esta transgresión –al igual que se intenta evitar la parcialidad judicial–, el tribunal que decide el destino del reo en el

¹⁰⁵ NIEVA FENOLL, Jordi. “Lar Razón De Ser De La Presunción De Inocencia”. *InDret*, vol.1, 2016, p. 15 y ss.

¹⁰⁶ *Ibid.*, p.10

proceso posterior no puede estar compuesto por aquél quien valoró el riesgo, a fin de no anticiparse a la pena. Podemos concluir que la presunción de inocencia es parcialmente compatible con la valoración del riesgo para la adopción de medidas cautelares.

Por el contrario, debemos señalar que, si esta valoración del riesgo –concretamente el riesgo de reiteración delictiva–, se realiza con otro fin que no sea para la adopción de medidas cautelares, estaríamos ante una vulneración ilegítima del derecho a la presunción de inocencia. Esto es así porque al intentar valorar este riesgo, se está suponiendo que el reo va a volver a delinquir sin que haya un motivo suficientemente fuerte que lo justifique. La expresión “es inocente hasta que se pruebe lo contrario” que define la presunción de inocencia carece de sentido cuando se intenta anticipar a la conducta delictiva dado que no puede ser probado algo que aún no ha sucedido.

Más aún se ve vulnerado este derecho cuando, para esta valoración, se hace uso de los algoritmos. Esto es así, porque las herramientas que utilizan inteligencia artificial no pueden predecir de forma cierta la conducta futura. Como ya señalamos anteriormente, estas no predicen la probabilidad en la que el reo vuelva a delinquir, no analiza la peligrosidad del individuo, sino que realiza una prognosis del riesgo a partir de los factores que se dan dentro de un contexto determinado. En este sentido, la valoración sería totalmente desindividualizada, teniendo en cuenta únicamente las características del reo que lo relacionan dentro de un determinado perfil de riesgo. En este sentido, la misma expresión a la que antes aludíamos –“es inocente hasta que se pruebe lo contrario”– también carece de sentido al hacer uso de estas herramientas puesto a que no se puede probar el actuar futuro de un individuo a partir de el actuar pasado de sus iguales.

Conclusiones

I. El objeto de este trabajo ha estado enfocado al estudio de la posibilidad de la utilización de herramientas que utilizan la inteligencia artificial para la valoración del *periculum libertatis*. Desde un comienzo, hemos podido observar que la valoración del *periculum* desde un punto de vista humano puede verse alterado por la presencia de heurísticos que dan lugar a una valoración del riesgo sesgada y a una falta de motivación en los autos que adoptan medidas cautelares. En primer lugar, el heurístico de disponibilidad, relacionado con la experiencia, traía consigo los sesgos de *optimismo* – infravalorar el riesgo– y *correlación ilusoria* –sobredimensionar el riesgo por realizar una correlación errónea–. Por otro lado, el heurístico de afecto –en relación con la presencia de emociones– podía dar lugar también a una valoración sesgada. Hemos visto que, tanto la ausencia como la intensa presencia de las emociones podía dar lugar a una incorrecta valoración del riesgo. Por ello, nos planteamos si la inteligencia artificial podía corregir estas dificultades en la valoración del riesgo y si podíamos confiar en ella con suficiente certeza como para limitar los derechos del reo mediante la adopción de medidas cautelares.

II. Para la valoración del riesgo de fuga y de destrucción de pruebas, hemos podido observar que tanto los presupuestos que se establecen en la ley como en la doctrina dan lugar a una objetivación de ambos riesgos, pudiendo ser viable la utilización de un algoritmo. Sin embargo, el problema que nos encontramos es que la concurrencia de estos presupuestos no implica necesariamente que el reo se fugue o destruya pruebas puesto que no se estaría prediciendo una conducta futura, sino que estaríamos ante el resultado de una probabilidad. Pese a ello, en la práctica sería indiferente que la valoración la hiciese un juez o un algoritmo, pues ambos tendrían en consideración los mismos presupuestos.

III. Por otro lado, respecto a la valoración del riesgo de reincidencia hemos podido apreciar ciertas dificultades que pueden comprometer los derechos del reo y de la víctima. En primer lugar, porque los algoritmos, a diferencia del ser humano, no pueden experimentar emociones, ni improvisar ante nuevas situaciones, por ello, es de vital importancia la intervención humana. Así lo vimos con el programa VioGén –epígrafe 3.1.3.1.–, que la omisión de las autoridades policiales y judiciales desencadenó un resultado fatal. Por ello, es inadmisibles la utilización de este tipo de programas de forma exclusiva y excluyente para la valoración del riesgo de reincidencia. Asimismo, pueden

estar llenos de prejuicios que reflejan los del propio programador al haber basado sus algoritmos en datos estadísticos recogidos de forma sesgada, como hemos podido comprobar en el estudio del programa COMPAS –epígrafe 3.1.3.2–, que resultó ser racista. Por otro lado, si el algoritmo aprendiese a partir de casos en los que la valoración del riesgo fue valorada de forma sesgada, daría lugar a una perpetuación y arraigo de estos sesgos.

IV. En cuanto a la necesidad de la intervención humana en este tipo de valoraciones, podemos encontrarnos también con el problema de que el decisor se confíe demasiado en el resultado que le proporciona la herramienta. Además, podría peligrar un mal uso hasta tal punto que, como hemos visto, se persiga y se acose a ciertos individuos por un delito que aún no ha sido cometido.

V. Otro de los problemas que se plantea es que la valoración que se realiza mediante los sistemas actuariales se hace sobre el riesgo, y no sobre la peligrosidad del individuo. No predicen una conducta delictiva, sino que determinan la probabilidad de que esta acontezca. No nos proporcionará un resultado causal, sino que será correlacional, dando lugar a una objetivación del riesgo. Estos sistemas establecen correlaciones entre casos que aparentemente podrían tener alguna conexión, pero en realidad, no hay forma de demostrar que ésta existe. Lo que se pretende es establecer patrones delictivos y aplicarlos la sociedad, produciéndose una situación en la que se valora a un individuo por las acciones de sus iguales y no por las suyas propias.

VI. La falta de transparencia que supone el uso de estas herramientas puede vulnerar el derecho de defensa establecido en el artículo 24 CE al imposibilitar al reo que despliegue su efectiva defensa frente a la valoración del riesgo determinada por el algoritmo.

VII. Asimismo, el derecho a la presunción de inocencia y el derecho a un juez imparcial, pueden verse afectados al intentar predecir la conducta delictiva. Como hemos podido observar, su vulneración en la valoración del *periculum* es inevitable. Por un lado, la transgresión del derecho a la presunción de inocencia es necesaria para poder preservar el objeto de la medida cautelar. Por otro lado, al establecer la necesidad de la valoración del riesgo de reincidencia para adoptar la medida cautelar, es necesario que el juez parta de la idea de que el reo puede volver a delinquir, posicionándose contra él. Estas vulneraciones, de algún modo, son necesarias en la fase de instrucción, y por ello, podríamos considerarlas legítimas. Lo que es desproporcionado e impensable es utilizar

los algoritmos para predecir la conducta delictiva dentro del proceso en el que se decida el destino del reo.

VI. Bajo mi punto de vista, no es recomendable la utilización de este tipo de herramientas para la valoración del *periculum*. En primer lugar, por los problemas inherentes a la utilización de los algoritmos en general, la presencia de los sesgos en los datos, ya sean introducidos por el programador o porque se base en casos que hayan sido resueltos de forma sesgada. Además, la presencia de estas herramientas implicaría necesariamente que el reo alegara su indefensión puesto que no es posible saber con certeza los factores en los que el algoritmo se ha basado para realizar la valoración del riesgo. Por otro lado, existe un alto riesgo de que el ser humano confíe ciegamente en ellas antes que en su propio juicio.

Fuentes y bibliografía

AGUILÓ, Josep. “Imparcialidad y concepciones del derecho”. *Juridicas*, vol. 6, n. 2, 2009, p. 27-44.

BERIAN, Iñigo de Miguel y PÉREZ, Miren Josune. “La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados”. *Revista de Derecho UNED*, n. 25, 2019, p. 531-561.

BORGES, Raquel. “El sesgo de la máquina en la toma de decisiones en el proceso penal”. *Ius et scientia*, vol. 6, n. 2, 2020, p. 54-71.

CATENA, Víctor Manuel. “Sobre el derecho de defensa: cuestiones generales”. *Teoría y Derecho*, n. 8, 2010 p. 17-40.

CHAPMAN, Loren J. “Illusory correlation in observational report”. *Journal of Verbal Learning and Verbal Behavior*, vol. 6, n.1, 1967, p. 151-155.

COLMENERO, Miguel. “La garantía del derecho a un juez imparcial”. *Persona y derecho*, vol. 55, 2006, p. 721-736.

DAMASIO, Antonio. *El error de Descartes. La emoción, la razón y el cerebro humano*. Barcelona, 2013.

DIETERICH, William, MENDOZA, Christina y BRENNAN, Tim. “COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity”. *Northpointe Inc. Research Department*, 2016, p. 1-39.

DRESSEL, Julia y FARID, Hany. “The accuracy, fairness, and limits of predicting recidivism”. *Science Advances*, vol. 4, n. 1, 2018, p. 1-5.

ESTÉVEZ, Lucana M. “Los derechos fundamentales ante las nuevas tecnologías: protección o vulneración a la luz de HART y VALCRI”, *Alternativas al sistema de justicia criminal latinoamericano*, p. 61-68.

ESTÉVEZ, Lucana. “Inteligencia artificial y violencia contra las mujeres: ¿funcionan los sistemas automatizados de evaluación del riesgo?” *PERSPECTIVAS Revista de ciencias jurídicas y políticas*, n. 3, 2020, p. 127-141.

FERGUSON, Andrew G. “Illuminating black data policing”. *Ohio St. J. Crim.*, vol. 15, n. 503, 2017, p. 503-525.

FERGUSON, Andrew G. “Policing predictive policing”. *Wash. UL Rev.*, vol. 94, n. 5, 2017, p. 1109-1189.

FINUCANE, Melissa, ALHAKAMI, Ali, SLOVIC, Paul, JOHNSON, Stephen. “The affect heuristic in judgments of risks and benefits”. *Journal of behavioral decision making*, vol. 13, n. 1, 2000, p. 1-17.

GONZÁLEZ-ÁLVAREZ, José Luis; Jorge SANTOS y Miguel CAMACHO, “Policía predictiva en España. Aplicación y retos futuros”. *Fundación Universitaria Behavior & Law*, vol. 6, n.1, 2020, p. 26-41.

GUTIÉRREZ, Juan David. “Retos éticos de la inteligencia artificial en el proceso judicial”, *XLI Congreso Colombiano de Derecho Procesal*, 2020, p. 1-23.

HECKER, Philipp. “Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law”. *Common Market Law review*, 2018, p.1-30.

HERNÁNDEZ, María. “Inteligencia artificial y Derecho Penal”, *Actualidad Jurídica Iberoamericana*, 10 bis, 2019, p. 792-843.

HIRSCHLAFF, Alan, BERMÚDEZ, Gastón, ROMERO, Juan, FRANCO, Micaela, STEFANOLI, Nicolás, ANTONINI, Santiago, TAGLIAMONTE, Santiago, GONZÁLEZ, Tomás. “Inteligencia artificial y el rol del Juez: ¿el fin de los jueces convencionales o la confirmación de su insustituibilidad?”. *XLI Congreso Colombiano del Derecho Procesal*, 2020, p. 1-33.

HOLZINGER, Andreas, PLASS, Markus, HOLZINGER, Katharina, CERASELA, Gloria, PINTEA, Camelia y PALADE, Vasile. “A glass-box interactive machine learning approach for solving NP-hard problems with the human-in-the-loop”, 2017, p. 1- 26.

JOFFE, Hélène. “Risk: From perception to social representation”. *British journal of social psychology*, vol. 42, n.1, 2003, p. 55-73.

JULIÀ, Miquel. “Proceso penal y (neuro)ciencia: una interacción desorientada”. Disponible en <http://diposit.ub.edu/dspace/handle/2445/143437>

KELLER, Carmen, SIEGRIST, Michale y GUTSCHER, Heinz. “The role of the affect and availability heuristics in risk communication”. *Risk Analysis*, 2006, vol. 26, n. 3, p. 631-639.

LARSON, Jeff. MATTU, Surya. KIRCHNER, Lauren y ANGWIN, Julia. “How We Analyzed the COMPAS Recidivism Algorithm”. *ProPublica*, 2016, p. 1-16.

LEGG, Michael. y BELL, Felicity. “Artificial Intelligence and the Legal Profession: A Primer”, *Flip stream*, 2020, p. 1-24.

MARTÍNEZ, Lucía. y MONTES, Francisco. “El uso de valoraciones del riesgo de violencia en derecho penal: algunas cautelas necesarias”, *InDret*, n.1, 2018.

MIRÓ, Fernando. “Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots”, *Revista de derecho penal y criminología*, vol.3, n. 20, 2018, p. 87-130.

MORA, Jeffry José. “Predictibilidad conductual y proceso penal: algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares”. *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, n. 2, 2021, p. 53-85.

NIEVA FENOLL, Jordi. *Derecho procesal III*. Marcial Pons, Madrid, 2017.

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

NIEVA FENOLL, Jordi. “La Razón De Ser De La Presunción De Inocencia”. *InDret*, vol.1, 2016.

NISSAN, Ephraim. “Digital technologies and artificial intelligence’s present and foreseeable impact on lawyering, judging, policing and law enforcement”. *Sprinter*, 2015, p. 441-464.

NOVO, Mercedes, ARCE, Ramón, FARIÑA, Francisca. “El heurístico: perspectiva histórica, concepto y tipología”. *Jueces: Formación de juicios y sentencias*, 2003, p. 39-66.

PACHUR, Thorsten., HERTWIG, Ralph, STEINMANN, Florian. “How do people judge risks: availability heuristic, affect heuristic, or both?”. *Journal of Experimental Psychology: Applied*, vol.18, n.3, p. 314-330.

PÁEZ, Andrés. “Los sesgos cognitivos y la legitimidad racional de las decisiones judiciales”. *Filosofía y psicología, Universidad Externado de Colombia*, 2020.

PUEYO, Antonio y REDONDO, Santiago. “Predicción de la violencia: entre la peligrosidad y la valoración del riesgo de violencia”. *Papeles del Psicólogo*, vol. 28, n.3, p. 157-173.

PUELLES, Luis; SÁNCHEZ, José y ALBERTOS Pedro, “Inteligencia Artificial e Inteligencia Humana”, *Sociedad Económica de Amigos del País*, 2010, p. 71-94.

ROMERO, Carlos María. “Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad”. *R.E.D.S.*, n.13, 2018, p. 39-55.

SHEEHEY, Bonnie. “Algorithmic paranoia: the temporal governmentality of predictive policing”. *Ethics and Information Technology*, vol21, n. 1, 2019, p. 49-58.

SLOVIC, Paul, FINUCANE, Melissa L., PETERS, Ellen y MACGREGOR, Donald. “Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality”. *Risk Analysis*, vol. 24, n. 2, p. 311-322.

SLOVIC, Paul, FISCHHOFF, Baruch y LICHTENSTEIN, Sarah. “The psychometric study of risk perception, *Springer, Boston, MA*, vol.1, 1986, p. 3-24.

SLOVIC, Paul, LICHTENSTEIN, Sarah, FIRSCHHOFF, Baruch. “Accident probabilities and seat belt usage: A psychological perspective”. *Accident Analysis & Prevention*, vol. 10, n. 4, 1978, p. 281-285.

SLOVIC, Paul, PETERS, Ellen, FINUCANE, Melissa, MACGREGOR, Donald. “Affect, risk, and decision making”. *Health psychology*, vol. 24, n. 4S, 2005.

STAJNOLOVIC, Milena. “Percepción social de riesgo: una mirada general y aplicación a la comunicación de salud”. *Revista de Comunicación y Salud*, vol. 5, 2015, p. 96-107.

TUCEK, Aron. “Constraining Big Brother: The Legal Deficiencies Surrounding Chicago's Use of the Strategic Subject List”. *U. Chi. Legal F*, v.2018, 2019, p. 427-460.

VAN DEN BRAAK, Susan W. y VREESWIJK, Gerard A. W. “A knowledge representation architecture for the construction of stories based on interpretation and evidence”, 2006.

WEINSTEIN, Neil D. “Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample”. *Journal of behavioral medicine*, vol. 10, n.5, 1987, p. 481-500.

Artículos de prensa

ÁLVAREZ BARCELONA, Ramón. “La inteligencia artificial de la Policía que desenmascara denuncias falsas”. *La Vanguardia*. 13 de abril de 2019 (disponible en <https://www.lavanguardia.com/tecnologia/20190414/461583468024/veripol-policia-nacional-inteligencia-artificial-algoritmo-denuncias-falsas.html>; última consulta 24/02/2021).

ASHER, Jeff y ARTHUR, Rob. “Inside the Algorithm That Tries to Predict Gun Violence in Chicago”. *The New York Times*, 13 de junio 2017 (disponible en <https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html>; última consulta 04/04/2021).

BURANYI, Stephen. “Rise of the racist robots – how AI is learning all our worst impulses”. *The Guardian*, 8 de agosto de 2017 (disponible en <https://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses>; última consulta 16/03/2021).

CAMACHO, Julia. “El fiscal pide que vuelva a prisión el guardia civil de “La manada” por riesgo de fuga”. *El periódico*, 28 de junio de 2018 (disponible en <https://www.elperiodico.com/es/sociedad/20180628/el-guardia-civil-de-la-manada-quiso-sacarse-un-pasaporte-al-quedar-en-libertad-6913116>; última consulta 14/03/2021).

FRESNEDA, Carlos. “Un ordenador logra superar por primera vez el test de Turing”. *El Mundo*, 19 de junio de 2014 (disponible en <https://www.elmundo.es/ciencia/2014/06/09/539589ee268e3e096c8b4584.html>; última consulta 20/02/2021).

PROTO, Lucas. “El ‘Minority report’ de Florida: un ‘sheriff’ persigue delitos todavía no cometidos”. *El Confidencial*, 15 de marzo de 2021 (disponible en https://www.elconfidencial.com/mundo/2021-03-15/el-minority-report-de-florida-un-sheriff-persigue-delitos-todavia-no-cometidos_2993108/; última consulta 05/03/2021).

REVELL, Timothy. “AI detective analyses police data to learn how to crack cases”. *Newscientist*, 10 de mayor de 2017 (disponible en <https://www.newscientist.com/article/mg23431254-000-ai-detective-analyses-police-data-to-learn-how-to-crack-cases/>; última consulta 01/03/2021).

Jurisprudencia

- **Tribunal Constitucional**

Sentencia del Tribunal Constitucional del 12 de julio 145/1988.

Sentencia del Tribunal Constitucional de 15 de abril 62/1996.

Sentencia del Tribunal Constitucional del 12 de junio 152/2000.

Sentencia del Tribunal Constitucional del 22 de marzo 39/2004.

Sentencia del Tribunal Constitucional del 12 de septiembre 215/2005.

Sentencia del Tribunal Constitucional del 26 de septiembre 233/2005.

Sentencia del Tribunal Constitucional del 3 de julio 219/2006.

- **Tribunal Supremo**

Sentencia del Tribunal Supremo Sala de lo Penal, Sección 1ª, de 17 de mayo 2018/2257.

- **Juzgado Central de Instrucción**

Auto Juzgado Central de Instrucción de 30 de octubre 2014/1197.

Auto Juzgado Central de Instrucción nº 6 del 3 de marzo 0000096/2017.

Auto Juzgado Central de Instrucción del 16 de octubre 2017/252463.

- **Juzgado de Instrucción**

Auto Juzgado de instrucción núm. 2 de Reus de 23 de marzo 2020/510.

- **Otros**

Sentencia del 13 de julio de 2016: State v. Loomis, 881, Wisconsin.

Legislación

Constitución española (BOE núm.311, de 29 de diciembre de 1978).

España. Real Decreto de 14 de septiembre de 1882 por el que se aprueba la Ley de Enjuiciamiento Criminal (BOE núm. 260, de 17 de noviembre de 1882).

España. Ley Orgánica 6/1985 de 1 de julio del Poder Judicial (BOE núm. 157 de 2 de julio de 1985).

España. Ley Orgánica 1/2004, de 28 de diciembre de Medidas de Protección Integral contra la Violencia de Género (BOE núm. 313, de 29 de diciembre de 2004).