

El uso de sistemas predictivos automatizables
en la actividad decisonal de las autoridades públicas:
aportaciones para un análisis iusfilosófico sobre
la digitalización del proceso decisonal

*The Use of Automatable Predictive Systems
in the Decision-Making Activity of Public Authorities:
Contributions to a Legal-Philosophical Analysis
of the Digitization of the Decision-Making Process*

Por ANTONIO MADRID PÉREZ
Universitat de Barcelona

RESUMEN

Este artículo se ocupa del impacto que el uso de los sistemas predictivos automatizables tiene, y puede tener en un futuro inmediato, en la actividad decisonal. Para abordar esta cuestión se han elegido tres aspectos que se consideran centrales: la inclusión de elementos evaluativos en los sistemas predictivos, el estatus normativo de los sistemas predictivos utilizados, y, en tercer lugar, la responsabilidad por la configuración formal y sustantiva del proceso decisonal.

ABSTRACT

This article studies the impact the use of automatable predictive systems has, and may have in the immediate future, on decision-making. To address this question, three aspects that are considered central have been chosen: the inclusion of evaluative elements in the predictive systems, the normative

status of the predictive systems used, and, thirdly, the responsibility for the formal and substantive configuration of the decision-making process.

SUMARIO: 1. PRESENTACIÓN DEL TEMA. – 2. LA UTILIZACIÓN DE SISTEMAS PREDICTIVOS EN LA TOMA DE DECISIONES: DE LA VALORACIÓN A LA PREDICCIÓN. 2.1 *Los sistemas predictivos como parte de la arquitectura decisional.* 2.2 *El uso de la IA en la mejora de la actividad decisional.* 2.3 *Entre la «variabilidad no deseada» y la necesidad de variabilidad.* – 3. EL ESTATUS NORMATIVO DE LOS SISTEMAS PREDICTIVOS UTILIZADOS POR LAS AUTORIDADES EN PROCESOS DECISIONALES. 3.1 *La prohibición de automatización de las decisiones individualizadas.* 3.2 *El derecho a saber y la prohibición de utilizar determinados tipos de datos en la automatización de decisiones.* – 4. A MODO DE CONCLUSIÓN: LA RESPONSABILIDAD POR LA CONFIGURACIÓN DE LA DECISIÓN.

SUMMARY: 1. PRESENTATION OF THE SUBJECT. – 2. THE USE OF PREDICTIVE SYSTEMS IN DECISION-MAKING: FROM ASSESSMENT TO PREDICTION. 2.1 *Predictive systems as part of the decisional architecture.* 2.2 *The use of AI in improving decision-making activity.* 2.3 *Between «unwanted variability» and the need for variability.* – 3. THE REGULATORY STATUS OF THE PREDICTIVE SYSTEMS USED BY THE AUTHORITIES IN DECISION-MAKING PROCESSES. 3.1 *The prohibition of automation of individualized decisions.* 3.2 *The right to know and the prohibition to use certain types of data in decision automation.* – 4. BY WAY OF CONCLUSION: RESPONSIBILITY FOR THE CONFIGURATION OF THE DECISION.

1. PRESENTACIÓN DEL TEMA

La inteligencia artificial (en adelante IA) se ha convertido en un objeto de estudio para la filosofía jurídica española¹, además de para

¹ Entre otros: CÁCERES NIETO, E., «La inteligencia artificial aplicada al Derecho como una nueva rama de la Teoría jurídica», *Anales de la Cátedra Francisco Suárez*, n. 57, 2023, pp. 63-89; MORONDO TARAMUNDI, D., «Discrimination by Machine-Based Decisions: Inputs and Limits of Anti-discrimination Law», en CUSTERS, B. y FOSCH-VILLARONGA, E., *Law and Artificial Intelligence. Regulating AI and Applying AI in Legal Practice*, The Hague, Asser Press, 2022, pp. 73-85; GARCÍA AÑÓN, M. J., «Desigualdades algorítmicas: conductas de alto riesgo para los derechos humanos», *Derechos y libertades*, n. 47, 2022, pp. 17-49; MARTÍNEZ DE PISÓN CAVERO, J., «Inteligencia artificial y los Derechos humanos. Viejos y nuevos retos de la Filosofía del Derecho», *Revista internacional de Pensamiento político*, I Época, vol. 17, 2022, pp. 377-387; Solar Cayón, J. I. y Sánchez Martínez, M.ª O. (dirs.), *El impacto de la inteligencia artificial en la teoría y la práctica jurídica*, Madrid, La Ley – Wolters Kluwer, 2022; Llano Alonso, F. H. (dir.), Martín Garrido, J. y Valdivia Jiménez, R. (coords.), *Inteligencia artificial y filosofía del derecho*, Murcia, Ediciones Laborum,

otras disciplinas jurídicas². Las tendencias actuales hacen pensar que IA va a acompañar y, al mismo tiempo, a condicionar progresivamente la decisión humana³. Pese a que nos hallamos en una fase histórica de optimismo tecnológico⁴ en la que a la IA se asocian expectativas que todavía no se han dado⁵, y hay quien duda de que algunas de las

2022; LLANO ALONSO, F. H., «La deshumanización del Derecho en la era digital», *Informática y Derecho. Revista Iberoamericana de Derecho Informático (segunda época)*, n. 12, 2022, pp. 91-108; MADRID PÉREZ, A., «La inteligencia artificial y la robótica como motores de cambio del Derecho», en Estévez Araujo, J. A. (ed.), *El Derecho ya no es lo que era. Las transformaciones jurídicas en la globalización neoliberal*, Madrid, Trotta, 2021, pp. 171-194; BONSIGNORE FOUQUET, D., «Sobre inteligencia argumental, decisiones judiciales y vacíos de argumentación», *Teoría y Derecho. Revista de pensamiento jurídico*, n. 29, 2021, pp. 248-279; RODRÍGUEZ PUERTO, M., «¿Puede la inteligencia artificial interpretar normas jurídicas? Un problema de razón práctica», *Cuadernos electrónicos de Filosofía del Derecho*, n. 44, 2021, pp. 74-96; CAMPIONE, R., *La plausibilidad del derecho en la era de la inteligencia artificial. Filosofía carbónica y filosofía silícica del derecho*, Madrid, Dykinson/Universidad Carlos III de Madrid, 2020; SOLAR CAYÓN, J. I., *La inteligencia artificial jurídica: el impacto de la innovación tecnológica en la práctica del derecho y el mercado de servicios jurídicos*, Madrid, Thomson Reuters-Aranzadi, 2019; MARTÍNEZ GARCÍA, J. I., «Derecho inteligente», *Cuadernos electrónicos de Filosofía del Derecho*, n. 37, 2018, pp. 95-114; ASÍS, R., de, *Una mirada a la robótica desde los derechos humanos*, Madrid, Dykinson, 2014; BELLOSO MARTÍN, N., «Entre la técnica y la ciencia del Derecho. ¿Hacia una hermenéutica telemática?», *Anales de la Cátedra Francisco Suárez*, n. 47, 2013, pp. 139-161; Bourcier, D. y Casanovas, P. (eds.), *Inteligencia artificial y derecho*, Barcelona, UOC, 2003; GALINDO AYUDA, F., *Derecho e Informática*, Madrid, La Ley, 1998.

² Entre otros: MEZA RIVAS, M., *Las armas autónomas letales: Un desafío para el Derecho internacional humanitario, los Derechos humanos, la Seguridad y el Desarme internacionales*, Valencia, Tirant lo Blanch, 2022; MERCADER UGUINA, J. R., *Algoritmos e inteligencia artificial en el derecho digital del trabajo*, Valencia, Tirant lo Blanch, 2022; PÉREZ DAUDÍ, V., *De la Justicia a la Ciberjusticia*, Barcelona, Atelier, 2022; Rivas Vallejo, M. P. (dir.), *Discriminación algorítmica en el ámbito laboral*, Thomson Reuters Aranzadi, 2022, GÓMEZ JIMÉNEZ, M. L., *Automatización procedimental y sesgo electrónico: el procedimiento administrativo electrónico desde la inteligencia artificial*, Madrid, Aranzadi, 2021; RAMIÓ MATAS, C., *Inteligencia artificial y administración pública. Robots y humanos compartiendo el servicio público*, Madrid, Los Libros de la Catarata, 2019; MOISÉS BARRIO, A., *Derecho de los Robots*, Madrid, Wolters Kluwer, 2018; NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial*, Barcelona, Marcial Pons, 2018.

³ Una visión crítica la ofrece, entre muchos otros, SADIN, É., *La silicolonización del mundo. La irresistible expansión del liberalismo digital*, Caja negra editora, Buenos Aires, 2018, p. 116. Para un breve repaso sobre la relación entre IA y Derecho, vid. RISSLAND, E. L., ASHLEY, K. D. y LOUI, R. P., «AI and Law: A fruitful synergy», *Artificial Intelligence*, n. 150, 2003, pp. 1-15.

⁴ Hay quien lo ha expresado como «idolatría de los datos y de la IA», vid. ADAMS BHATTI, S., «Algorithms in the Justice System: Current Practices, Legal and Ethical Challenges», en MATT HERVEY, M. y LAVY, M., *The Law of Artificial Intelligence*, Thomson Reuters, 2021, pp. 523-555, p. 525 ss.

⁵ Hay que atender al ciclo de sobre-expectación (*Gartner Hype Cycle*) acuñado por Gartner Group (empresa especializada en las tecnologías de la información) a finales de los años 90. Se refiere a las expectativas depositadas en las nuevas tecnologías cuando son lanzadas y a cómo estas expectativas muchas veces no quedan satisfechas. De forma complementaria, también conviene tener

expectativas creadas se puedan dar⁶, lo cierto es que los sistemas algorítmicos crecientemente forman parte de procesos decisionales⁷. Ente las principales preocupaciones que plantea el uso de estos mecanismos está que los sistemas que inicialmente han sido pensados como herramientas de apoyo a la toma de decisiones acaben sustituyendo la decisión humana⁸.

Para especialistas contemporáneos como Russell y Norvig es habitual decir que un sistema decide. Hablan en este sentido de agentes capaces de tomar decisiones, entendiendo por tal: «aquél capaz de tomar decisiones racionales basándose en lo que cree y desea»⁹. Para estos autores, el agente «decide» si es capaz de hacerlo en contextos de incertidumbre y objetivos contradictorios¹⁰. De esta forma, un sistema autónomo de conducción podría ser programado y entrenado para «tomar decisiones» en situaciones de incertidumbre. Sin embargo, para otros autores, la capacidad racional de decisión exige que el agente pueda modificar los objetivos de su acción a partir de sus valo-

presente la conocida como ley de Amara (por Roy Amara): la tendencia a sobreestimar los efectos de una tecnología en el corto plazo y subestimar el efecto en el largo plazo.

⁶ LARSON, E. J., *El mito de la IA: por qué las máquinas no pueden pensar como nosotros lo hacemos*, Bilbao, Shackleton Books, 2022; SEARLE, J., «Can Computers Think?», en Chalmers, D. J. (ed), *Philosophy of Mind: Classical and Contemporary Readings*, New York, Oxford University Press, 2002, pp. 669-675.

⁷ Desde 2017, China ha habilitado tres tribunales telemáticos para procesar y dictar sentencias sobre materias civiles y administrativas vinculadas al uso de internet en las circunscripciones de Hangzhou, Beijing y Guangzhou. *Vid.*, en un elogio de este sistema, GUO, M., «Internet' courts challenges and future in China», *Computer Law and Security Review*, n. 40, 2021, pp. 1-13 <https://doi.org/10.1016/j.clsr.2020.105522> Para una crítica de este tipo de tribunales desde un modelo garantista, *vid.* SUNG, H., ¿«Can Online Courts Promote the Access to Justice? A case study of the Internet Courts in China», *Computer Law and Security Review*, n. 39, 2020, pp. 1-15, <https://doi.org/10.1016/j.clsr.2020.105461> y XU, Z., ZHAO, Y. y DENG, ZHONGWEN, «The possibilities and limits of AI Chinese judicial judgment», en *AI & Society*, 2021, pp. 1-11, <https://doi.org/10.1007/s00146-021-01250-9>

Vid. tb. SUSSKIND, R., *Online Courts and the Future of Justice*, Oxford, 2019; Susskind, R. y SUSSKIND, D., *El futuro de las profesiones. Cómo la tecnología transformará el trabajo de los expertos humanos*, TEELL Editorial, 2016, pp. 65-69 y BARONA VILAR, S., *Algoritmización del Derecho y de la Justicia. De la inteligencia artificial a la Smart Justice*, Valencia, Tirant Lo Blanch, 2021, en especial pp. 365 ss., 580 ss., 625 ss.

⁸ *Vid.* BARONA VILAR, S., *op. cit.*, p. 560, 640 ss.; TODD, P. y BENSABAT, I., «The Influence of Decision Aids in Choice Strategies: An Experimental Analysis of the Role of Cognitive Effort», *Organizational Behavior and Human Decision Processes*, n. 60 (1), 1994, pp. 36-74; SOURDIN, T., «Judge v Robot? Artificial Intelligence and Judicial Decision-Making», *UNSW Law Journal*, n. 41(4), 2018, pp. 1114-1133, p. 1123.

⁹ RUSELL, S. J. y NORVIG, P., *Inteligencia artificial. Un enfoque moderno*, Madrid, Pearson / Prentice Hall, 2004, p. 663.

¹⁰ PARKES, D. C. y WELLMAN, M. P., «Economic reasoning and artificial intelligence», *Science*, vol. 349, 2015, pp. 267-72, p. 267.

res¹¹. En este último sentido, todavía no se ha desarrollado este tipo de racionalidad artificial.

Entre los sistemas utilizados en procesos decisionales se hallan aquellos que son diseñados y entrenados con finalidades predictivas¹². Este artículo aborda el impacto que el uso de estos sistemas predictivos tiene, y puede tener en un futuro inmediato, en la actividad decisional. Para abordar esta cuestión se han elegido tres aspectos que se consideran centrales: el papel que las predicciones emitidas por estos sistemas pueden desempeñar en el proceso decisional, con especial atención a la inclusión de elementos evaluativos en los sistemas predictivos, el estatus normativo de los sistemas predictivos utilizados, y, en tercer lugar, una reflexión sobre la responsabilidad no tanto por la decisión tomada, como por la configuración del proceso decisional que lleva a la toma de la decisión¹³.

En la filosofía jurídica española, los estudios sobre el uso de la IA en procesos decisionales tienen su antecedente más directo en el trabajo desarrollado por Antonio Enrique Pérez Luño en los años 70 del siglo pasado. En aquel momento, Pérez Luño ya abordó la relación entre el razonamiento jurídico y el razonamiento cibernético¹⁴. Planteó la posibilidad de un razonamiento cibernético instrumental, «un pensamiento ejecutivo lógico-mecánico»¹⁵ no sustitutivo respecto del razonamiento jurisprudencial¹⁶. Para Pérez Luño, en aquellos años, no se estaba ante la «automación del proceso decisorio» ya que la tecnología disponible no ofrecía esta posibilidad¹⁷. 20 años después, en los años 90, mantenía su preocupación sobre cuestiones centradas en la automatización del proceso decisorio. Por ejemplo, el riesgo que

¹¹ Vid. TOTSCHNIG, W., «Fully Autonomous AI», *Science and Engineering Ethics*, n. 26, 2020, pp. 2473–2485 <https://doi.org/10.1007/s11948-020-00243-z>; y la revisión crítica, que sobre este texto hace DENNIS, L. A., «Computational Goals, Values and Decision-Making», *Science and Engineering Ethics*, n. 26, 2020, pp. 2487-2495.

¹² Para seguir las implicaciones prácticas de estos sistemas, vid. MARTÍNEZ GARAY, L., «La incertidumbre de los pronósticos de peligrosidad: consecuencias para la dogmática de las medidas de seguridad», *Indret: Revista para el Análisis del Derecho*, 2(2), 2014, pp. 1-77, p. 15 ss.

¹³ No se aborda en este texto el uso de sistemas que tratan de predecir qué decisiones tomarán los tribunales, para esta cuestión, como introducción: MEDVEDEVA, M., VOLS, M. y WIELING, M., «Using machine learning to predict decisions of the European Court of Human Rights», *Artificial Intelligence and Law*, n. 28, 2019, pp. 237-266.

¹⁴ PÉREZ LUÑO, A. E., *Cibernética, Informática y Derecho (Un análisis metodológico)*, Bolonia, Studia Albornotiana XXVIII, Real Colegio de España, 1976.

En la década de los 70, en EE. UU., M. E. FRANKEL apostaba decididamente por la utilización de computadoras como un instrumento que aseguraba la aplicación de reglas en la actividad decisional de los tribunales: *Criminal Sentences: Law Without Order*, New York, Hill and Wang, 1972.

¹⁵ PÉREZ LUÑO, A. E., *op. cit.* p. 111.

¹⁶ *Idem*, p. 83 ss.

¹⁷ *Idem*, p. 111.

suponía potenciar el formalismo y la rigidez del procedimiento administrativo mediante la lógica «infalible, inexorable e inapelable de la programación informática»¹⁸. Sobre la actividad judicial, señalaba la limitación de los sistemas informáticos para comprender la multiplicidad de circunstancias que concurren en las conductas humanas. A su vez, hacía énfasis en un aspecto que sigue siendo central hoy en día, como es que el juez no debe delegar en un sistema informático su responsabilidad por la decisión que toma: «en la actualidad no es posible, ni deseable, una suplantación plena del razonamiento jurídico del juez o del abogado por el cálculo informático del ordenador. Sólo en aspectos de la experiencia jurídica rutinarios, estandarizados, formalizables, con variables predeterminadas cerradas es posible recurrir a sistemas expertos capaces de ofrecer soluciones operativas. Pero incluso en esos casos el juez no puede abdicar de la responsabilidad de su decisión para delegarla en el ordenador»¹⁹. Pasado el tiempo, estas observaciones siguen siendo pertinentes ya que el ámbito de reflexión y de preocupación acerca del uso de la IA en los procesos decisionales se ha expandido en la medida en que se han incrementado las prestaciones que ofrecen los sistemas algorítmicos.

2. LA UTILIZACIÓN DE SISTEMAS PREDICTIVOS EN LA TOMA DE DECISIONES: DE LA VALORACIÓN EN LA PREDICCIÓN

Aunque sea obvio lo que se va a decir a continuación, es preciso recordar que los sistemas de IA son creaciones humanas²⁰ que incorporamos en los procesos decisionales. Se ha de evitar el riesgo de anonimizar los sistemas algorítmicos. Estos sistemas no aparecen por generación espontánea, ni se utilizan, por azar. El uso de estos sistemas responde a una voluntad de las autoridades competentes, voluntad que puede ser explicitada con mayor o menor claridad.

Se han establecido distintas clasificaciones de las funciones que han de cumplir y que realmente cumplen los sistemas de IA en los procesos decisionales. Siguiendo a Herbert Simon, bajo el denominador común de *Contextual factors of algorithmic system*, se han descrito seis tipos de funciones: *Minimal automation*, *Acquisition and presentation of data*, *Suggested procedure steps*, *Supported decisions*, *Automated decisions* y *Autonomous decisions*²¹. Esta clasificación, al

¹⁸ PÉREZ LUÑO, A. E., *Manual de informática y derecho*, Barcelona, Ariel, 1996, pp. 85-6.

¹⁹ PÉREZ LUÑO, A. E., *Manual de informática y derecho*, op. cit., p. 97.

²⁰ FLORIDI, L., *Ethics, Governance, and Policies in Artificial Intelligence*, Springer, 2021, pp. 29-30.

²¹ ROEHL, U. B. U., «Understanding Automated Decision-Making in the Public Sector: A Classification of Automated, Administrative Decision-Making», en Juell-Skielse, G., Lindgren, I., Åkesson, M. (eds.), *Service Automation in the Public Sec-*

igual que otras disponibles²², es meramente orientativa. Como se verá en los sistemas analizados en este texto, si el uso de los sistemas se estudia desde la perspectiva de la configuración del proceso decisonal, se asiste a una combinación de funciones más que a una segmentación clara de las funciones prestadas por los sistemas.

Los sistemas predictivos, tanto en las funciones inicialmente pretendidas como en las realmente alcanzadas, pertenecen al contexto institucional en el que son utilizados²³. Son utilizados en ámbitos decisionales como la previsión de la posibilidad de reincidencia delictiva, los servicios bancarios, las inversiones bursátiles, el cálculo de riesgos, la policía predictiva o en otros ámbitos en los que se trata de predecir la conducta humana²⁴.

2.1 Los sistemas predictivos como parte de la arquitectura decisonal

Los sistemas con funciones predictivas toman la forma de sistemas expertos y de sistemas de aprendizaje automático. Los primeros se basan en el conocimiento de profesionales y especialistas a partir del cual se codifican en lenguaje computacional las reglas que son utilizadas para tomar decisiones sobre cuestiones legalmente acotadas: «sistemas de logros técnicos o de experiencia profesional que organizan grandes áreas de entorno material y social en el que vivimos»²⁵. Los segundos sistemas se basan en un proceso estadístico-probabilístico

tor. Progress in IS, Springer, 2022, pp. 35-63, p. 42 ss. A este respecto véase los comentarios de MORAL SORIANO, L., «Decisiones automatizadas, derecho administrativo y argumentación jurídica», en Llano Alonso, F. H. (dir.), Garrido Martín, J. y Valdivia Jiménez, R. (coords.), *Inteligencia artificial y filosofía del derecho*, Murcia, Ediciones Laborum, 2022, pp. 475-500, p. 480 ss.

²² SOLAR CAYÓN, J. I., «Inteligencia artificial y Justicia digital», en Llano Alonso, F. H. (dir.), Garrido Martín, J. y Valdivia Jiménez, R. (coords.), *Inteligencia artificial y filosofía del derecho*, Murcia, Ediciones Laborum, 2022, pp. 381-427, p. 387 ss; SHERIDAN, T. B., *Telerobotics, automation and human supervisory control*, MIT Press, 1992; Parasuraman, R., SHERIDAN, T. B. y WICKENS, C. D., «A model for types and levels of human interaction with automation», *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, n. 30 (3), 2000, pp. 286-297. <https://doi.org/10.1109/3468.844354>

²³ AGUIAR, F., «Teoría de la decisión e incertidumbre: modelos normativos y prescriptivos», *Empiria. Revista de metodología de las ciencias sociales*, n. 8, 2004, pp. 139-160, p. 142 ss.

²⁴ Sobre la evaluación de la fiabilidad de la máquina comparada con el juicio humano, *vid.* YOUYOU, W., KOSINSKI, M. y STILL WELL, D., «Computer-based personality judgments are more accurate than those made by humans», *PNAS. Proceedings of the National Academy of Sciences*, vol. 112, n. 4., 2015, pp. 1036-1040 www.pnas.org/cgi/doi/10.1073/pnas.1418680112. *Vid. tb.* WU, T., «Will artificial intelligence eat the Law? The rise of hybrid social-ordering systems», *Columbia Law Review*, vol. 119, n. 7, 2019, pp. 2001-2028.

²⁵ Véase cómo abordaban autores como GIDDENS, A., esta cuestión hace más de tres décadas: *Consecuencias de la modernidad*, Madrid, Alianza, 1993, p. 37.

del que se deriva una o varias reglas o procedimientos mediante los que se explican hechos o se predicen hechos futuros. En este segundo sentido, un modelo o sistema predictivo es *mathematical device for estimating unknown «variables» from known «variables»*, entendiéndose por «variable» *any quantity corresponding to some measurable feature of the world*²⁶. Es decir, es una forma de intentar conocer datos que no conocemos. En el ámbito jurídico, y a efectos de lo que interesa tratar en este texto, conocer datos que no conocemos tiene relevancia en la medida en que estos datos sean utilizados para calcular la probabilidad con la que se puede dar un hecho futuro.

Los sistemas predictivos y los usos que de ellos se hace en los procesos decisionales forman parte de lo que aquí vamos a denominar como «arquitectura decisional o arquitectura de la decisión»²⁷. La actividad decisional puede ser analizada como la sucesión de decisiones conectadas sin las cuales no sería posible la decisión final. Como ejemplo de análisis, se propone un proceso decisional regulado legalmente en el que distintas autoridades utilizan un sistema predictivo como apoyo a la competencia decisional que la ley les atribuye.

En 2019, la Audiencia provincial de Barcelona tuvo que decidir sobre el siguiente caso: la Junta de tratamiento del Centro penitenciario de Quatre Camins acordó conceder un permiso ordinario de salida de tres días a favor de un reo (decisión 1). El Juzgado de vigilancia penitenciaria competente autorizó el permiso (decisión 2). El Ministerio fiscal recurrió el auto del Juzgado de Vigilancia Penitenciaria (decisión 3). La defensa legal del reo presentó escrito oponiéndose al recurso de la Fiscalía (decisión 4). El recurso de la Fiscalía fue admitido a trámite por la Audiencia Provincial (decisión 5). Este tribunal concedió el permiso (decisión 6).

La Fiscalía fundamentó su recurso en la evaluación del riesgo de reincidencia que se había realizado previamente mediante el sistema que contiene la escala de valoración RisCanvi²⁸. La Junta de tratamiento y el Equipo técnico que avalaban la concesión del permiso eran conocedores del informe que había emitido el sistema algorítmico, según el cual el solicitante tenía un riesgo medio de reincidencia. La Fiscalía argumentó que con este informe no era posible conceder

²⁶ ZERILLI, J. y WELLER, A., «The Technology», en HERVEY, M. y LAVY, M., *The Law of Artificial Intelligence*, Thomson Reuters, 2021, pp. 7-30, pp. 15-6.

²⁷ La expresión «arquitectura decisional» es de uso habitual en informática. También se utiliza, por ejemplo, en la economía conductual. En los últimos años, R. H. THALER ha utilizado esta categoría de forma habitual para referirse a los factores que inciden en la toma de una decisión. *Vid.*, por ejemplo, *Todo lo que he aprendido con la psicología económica*, Bilbao, Deusto, 2016 y el texto publicado junto con C. R. SUNSTEIN, *Un pequeño empujón: el impulso que necesitas para tomar mejores decisiones sobre salud, dinero y felicidad*, Madrid, Taurus, 2017.

²⁸ *Vid.* «Manual de aplicación del protocolo de valoración RisCanvi», Subdirección general de Programas de Rehabilitación y Sanidad, Generalitat de Catalunya, 2019; MAGRO SERVET, V., «La inteligencia artificial para mejorar la lucha contra la violencia de género», *La ley*, 23 julio de 2021.

un permiso de salida. Por su parte, la Audiencia provincial tuvo en cuenta el resultado emitido por el RisCanvi, si bien argumentó en estos términos: «tampoco debemos obviar el resto de las circunstancias que afectan a la decisión. Hay que referirse además a la verdadera incidencia que los protocolos de evaluación y gestión han de tener en la decisión judicial que, al margen del carácter meramente orientador y no vinculante que han de tener sus resultados, habrá de ponderar la totalidad de los factores concurrentes y los informes de los distintos profesionales para llegar a un pronóstico de peligrosidad individualizado, huyendo de cualquier decisión automatizada en uno u otro sentido, destacando el auto recurrido que se mantienen riesgos bajos de reincidencia y quebrantamiento de condena según la valoración conjunta del psicólogo y el jurista»²⁹.

En este caso, la Audiencia tomó en consideración el resultado emitido por el sistema RisCanvi³⁰ pero decidió apartarse del parecer emitido por el sistema³¹. Para ello eligió otros elementos predictivos en los que fundamentar su decisión y, de forma explícita, argumentó en contra del criterio que ofrecía el sistema algorítmico. Bien podría haber seguido otro camino argumentativo. Podría haber fundamentado su decisión en el criterio emitido por el sistema, como previamente había hecho la Fiscalía al fundamentar su recurso. De hecho, desconocemos en cuántos casos el resultado que arroja este sistema es aceptado por las distintas autoridades implicadas en la concesión o denegación de un permiso carcelario.

A diferencia de sistemas como los que se utilizan en un gestor de expedientes para organizar la tramitación de documentos o para ordenar un procedimiento informativo, la característica de sistemas como el RisCanvi no es solo su funcionalidad predictiva³², sino, visto desde

²⁹ Considerando segundo del Auto de la Audiencia Provincial de Barcelona 938/2019, de 30 de mayo, ROJ: AAP B 5620/2019 –ECLI: ES: APB:2019:5620A; otros Autos concordantes en los que se clarifica «el carácter meramente orientador y no vinculante» de los resultados arrojados por sistemas como el RisCanvi, también identificados como «protocolos de evaluación y gestión» o como «guías de valoración del riesgo de comportamiento de abuso y agresión sexual» como el SVR-20: Auto de la Audiencia Provincial de Barcelona 339/2019, de 11 de febrero, ROJ: AAP B 1703/2019 –ECLI: ES: APB:2019:1703A; Auto de la Audiencia Provincial de Barcelona 403/2019, de 28 de febrero, ROJ: AAP B 3023/2019 –ECLI: ES: APB:2019:3023.^a, entre otros concordantes.

³⁰ El diseño y funcionamiento del RisCanvi sería un ejemplo de atrincheramiento de generalizaciones probabilísticas que se utilizan como generalizaciones universales, *vid.* SCHAUER, F., *Las reglas en juego. Un examen filosófico de la toma de decisiones basadas en reglas en el derecho y en la vida cotidiana*, Barcelona, Marcial Pons, 2004, p. 105.

³¹ MILLAR, J. y KERR, I., «Delegation, relinquishment, and responsibility: The prospect of expert robots», en CALO, R., FROMKIN, M. y KERR, I., *Robot Law*, Cheltenham, Edward Elgar, 2016, pp. 103-127, p. 119 ss.

³² Para un análisis crítico acerca de la capacidad predictiva del sistema RisCanvi, *vid.* MARTÍNEZ GARAY, L., «Errores conceptuales en la estimación de riesgo de reincidencia. La importancia de diferenciar sensibilidad y valor predictivo, y estimaciones

la perspectiva de la toma de decisiones, su inserción en el proceso que conduce a la toma de una decisión jurídicamente vinculante.

2.2 El uso de la IA en la mejora de la actividad decisional

La creciente utilización de los sistemas de IA en procesos decisionales ha intensificado el debate en torno a la mejora del juicio humano mediante el uso de la IA. Sin olvidar otras razones que impulsan el uso de la IA en procesos decisionales como la celeridad de computación, el manejo de grandes cantidades de datos, la capacidad para encontrar correlaciones o el ahorro de costes, la posible mejora del juicio humano mediante el uso de sistemas algorítmicos es una cuestión de enorme calado.

Los sistemas predictivos se basan en la demostrabilidad de las causas que han provocado supuestos que ya han tenido lugar. Como tales, un sistema predictivo expresa una cognición probabilística sobre hechos³³, no sobre normas. Nuestra confianza en la fiabilidad de los sistemas de predicción se basa en los aciertos que alcanzan las predicciones una vez podemos comprobar en qué medida se cumple la predicción. Por este motivo, en ámbitos como la predicción meteorológica, la mejora en la comprensión científica de los factores que determinan los fenómenos atmosféricos ha permitido mejorar la fiabilidad de los sistemas de predicción. En el ámbito del transporte, si hemos de realizar un viaje, nos fiamos de las indicaciones que nos proporcionan aplicaciones como Google Maps porque en nuestra experiencia este sistema nos dirige correctamente a nuestro destino, o nos informa correctamente acerca de cuál es el trayecto más rápido para llegar a un destino. Confiamos en la fiabilidad de estos sistemas³⁴, pero, crecientemente, tanto por el cambio de nuestros hábitos y prácticas sociales y profesionales, como por la propia capacidad de las aplicaciones para articularse con otras aplicaciones, cabe preguntarse si se amplía o se reduce nuestro margen de elección. Se producen dos fenómenos que pueden incidir directamente en la actividad de los decisores: en primer

de riesgo absolutas y relativas», *Revista Española de Investigación Criminológica*, 14 (3), 2016, pp. 11-31. La respuesta a esta crítica acerca de la capacidad predictiva de este sistema ha sido dada recientemente en Capdevila Capdevila, M. (coord.), *Las reincidencias en las excarcelaciones de alto riesgo (2014-16)*, Centro de estudios jurídicos y formación especializada, Generalitat de Catalunya, 2022.

³³ Vid. LAPORTA, F. J., *El imperio de la ley. Una visión actual*, Madrid, Trotta, 2007, p. 138 ss., especialmente la distinción entre la predicción de hechos futuros e inferencias lógicas o fácticas derivadas de hechos o razonamientos precedentes, a las que se refiere como «corolarios».

³⁴ LOGG, J. M., MINSON, J. A. y MORRE, D. A., «Algorithm Appreciation. People prefer Algorithmic to Human Judgment», *Organizational Behavior and Human Decision Processes*, vol. 151, 2018, pp. 90-103; SLONIM, N. et al., «An autonomous debating system», *Nature*, n. 591, 2021, pp. 379-384, p. 380, <https://doi.org/10.1038/s41586-021-03215-w>; Barona Vilar, Silvia, *op. cit.*, p. 555 ss.

lugar, que para alcanzar determinados objetivos se haga preciso utilizar estos sistemas; en segundo lugar, que en la medida en que estos sistemas se convierten en instrumentos insertos en las prácticas sociales y profesionales, pueden generar un sentido común que homogeneice las previsiones profesionales sobre hechos futuros³⁵. Estos sistemas van más allá de una mera tabla de cálculo de probabilidades, ya que generan patrones de interpretación³⁶.

Autores como Kleinberg argumentan que los sistemas algorítmicos pueden mejorar la toma de decisiones. Sin embargo, son prudentes: disponer de buenos predictores no garantiza tomar buenas decisiones. Especialmente si se adopta una mirada amplia sobre los impactos de las decisiones y sobre los objetivos que queremos que persigan estas decisiones. Existe una gran distancia entre construir un sistema algorítmico y las decisiones que se tomen con base en los resultados aportados por este sistema³⁷.

Menos precavidos han sido otros autores como Daniel Kahneman, Olivier Sibony y Cass R. Sunstein, que, en *Ruido. Un fallo en el juicio humano*³⁸, han defendido la mejora de la actividad decisional mediante el uso de la IA. Estos autores defienden la siguiente tesis: los sistemas de IA pueden reducir los errores en el juicio humano. Para desarrollar su estudio se han basado las decisiones de tribunales penales, las decisiones de compañías de seguros y los diagnósticos médicos. Para estos autores, la decisión humana se ve perjudicada por dos errores básicos: el sesgo (entendido como desviación sistémica en la toma de decisiones) y el ruido (entendido como dispersión aleatoria en la toma de decisiones³⁹). Tanto el sesgo como el ruido serían variaciones

³⁵ Vid. SANTOS, B. de S., *Crítica de la razón indolente. Contra el desperdicio de la experiencia. Vol. I. Para un nuevo sentido común: la ciencia, el derecho y la política en la transición paradigmática*, Bilbao, Desclée de Brouwer, 2000, p. 374 ss. Pese a la distancia temporal, tiene interés preguntarse cómo abordarían autores clásicos las novedades que ha introducido el uso de sistemas predictivos automatizables. En este sentido, cabría preguntarse, por lo menos especulativamente, si estas herramientas algorítmicas pueden actuar como fuentes del derecho en la medida en que, en tanto que elementos fácticos incorporados al sistema jurídico, establecen recomendaciones orientadas a la toma de decisiones jurídicamente vinculantes, vid. una posible aplicación desde la teorización de ROSS, A., *Teoría de las fuentes del Derecho*, Madrid, Centro de Estudios Político y Constitucionales, 2007, pp. 374-377.

³⁶ Estos patrones podrían llegar a funcionar en el proceso decisional como estándares que utilizara el juez u otra autoridad en la toma de decisiones. Para la noción de estándar de la que hago uso, vid. RUIZ MANERO, J., *El legado del positivismo jurídico. Ocho ensayos sobre cinco autores positivistas: Hans Kelsen, Norberto Bobbio, Eugenio Bulygin, Luigi Ferrajoli, Riccardo Guastini*, Lima, Palestra, 2014, p. 208 ss.

³⁷ KLEINBERG, J. et al., «Human Decisions and Machine Predictions», *The Quarterly Journal of Economics*, vol. 133, n. 1, 2018, pp. 237-293, pp. 243-44 y p. 288, doi:10.1093/qje/qjx032

³⁸ Madrid, Debate, 2021.

³⁹ Esta preocupación coincide con la intensificación en el uso de los instrumentos de evaluación de riesgos (*risk assessment instruments*). Uno de los objetivos que se buscaba al aplicar estos instrumentos era eliminar o minimizar los elementos subjetivos en la toma de decisiones. Acerca de la evolución de esta corriente en el pensa-

no deseadas («variabilidad no deseada») en juicios que idealmente deberían tener un mismo resultado⁴⁰. La propuesta que presentan es utilizar los algoritmos para reducir la variabilidad no deseada mediante la formalización de las decisiones. Los algoritmos, dicen, podrían hacerlo, tanto porque mejoran la capacidad de predicción, como por su «ausencia de ruido»⁴¹.

Para estos autores la IA supera en capacidad predictiva al juicio humano. Y es sobre esta funcionalidad predictiva que centran su análisis y sus propuestas no exentas, a mi parecer, de exclusiones argumentativas que restan consistencia a su posicionamiento⁴². Para defender su tesis central, distinguen entre dos tipos de juicios: juicios predictivos y juicios evaluativos⁴³. Reconocen que las decisiones se forman a partir de la combinación de juicios predictivos y de juicios evaluativos⁴⁴: «Condenar a un malhechor no es una predicción. Es un “juicio evaluativo” que busca adecuar la pena a la gravedad de delito»⁴⁵. Sin embargo, al elegir contratar a una persona entre varios candidatos, se combinan juicios predictivos (sobre su rendimiento futuro, enfermedades o situaciones personales de baja, cómo trabajará en equipo, qué liderazgo desarrollará...) con juicios evaluativos: las decisiones finales implican un equilibrio entre pros y contras de las opciones disponibles y «estos equilibrios se consiguen por medio de juicios evaluativos»⁴⁶.

Esta defensa tanto de la capacidad predictiva como de su uso en los procesos decisionales no es compartida por autores como Michael Tonry, uno de los principales penalistas y criminólogos experto en la materia, que ha analizado críticamente el uso de predicciones en la actividad decisional de los tribunales. Tonry señala problemas que deberían impedir el uso de predicciones en la formulación de sentencias: la escasa precisión de los sistemas de IA en la predicción de

miento criminológico, *vid.* HARCOURT, B. E., *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*, Chicago, University of Chicago Press, 2007; BRANDÁRIZ GARCÍA, J. Á., *El modelo gerencial-actuarial de penalidad. Eficiencia, riesgo y sistema penal*, Dykinson, Madrid, 2016.

⁴⁰ KAHNEMAN, D., SIBONY, O. y SUNSTEIN, C. R., *Ruido. Un fallo en el juicio humano*, Madrid, Debate, 2021, pp. 39, 49.

⁴¹ KAHNEMAN *et al.*, *op. cit.*, pp. 18 y 45.

⁴² Por una parte, afirman la superioridad de los sistemas de IA, pero por otra parte alertan contra el exceso de confianza en la predicción automatizada, KAHNEMAN *et al.*, *op. cit.*, pp. 160-2. Desconfían de la deliberación como forma de toma de decisiones por considerar que incrementan el ruido en la toma de decisiones: «... los grupos que deliberan tienden a ser más ruidosos que los grupos estadísticos, que se limitan a promediar los juicios individuales», p. 121. No tienen en cuenta la posibilidad de establecer reglas de deliberación y, por tanto, no contemplan la ponderación: Cfr. OTTOLIA, A., *Derecho, Big Data e inteligencia artificial*, Valencia, G. Giappichelli Editore-Tirant lo Blanch, 2018.

⁴³ KAHNEMAN *et al.*, *op. cit.*, pp. 59-69.

⁴⁴ *Idem*, p. 69.

⁴⁵ *Idem*, pp. 66-67.

⁴⁶ *Ibidem*.

violencias futuras⁴⁷, o el tipo de variables utilizadas al diseñar y entrenar el sistema que pueden ser tratadas como elecciones erróneas o sospechosas: la raza, la edad, el sexo o las opciones de vida como puede ser vivir en pareja, contraer matrimonio o vivir en un determinado barrio.

Al diseñar un sistema predictivo se parte de un número limitado de factores que se considera que pueden incidir en el hecho futuro. La misma selección de los factores responde a una evaluación previa realizada sobre la incidencia de los distintos factores. Si se considerara que un factor no guarda relación alguna con el hecho que se intenta predecir, o que su impacto es residual, este factor no aparecería en el listado de factores seleccionados. De esta forma, se puede decir que los sistemas predictivos contienen y reproducen una visión de la genealogía del hecho a predecir. En la medida en que el sistema predictivo establece una probabilidad de que el hecho suceda, y en la medida en que el sistema predictivo se construye a partir de la evaluación de los factores que pueden explicar que suceda el hecho futuro, los sistemas pueden ser analizados como un resultado de un proceso de teorización sobre el hecho futuro a partir de la teorización realizada sobre hechos pasados.

En el ámbito jurídico, el papel que pueden desempeñar las predicciones es diferente al que cumple una predicción en el ámbito de la meteorología ya que en el ámbito jurídico decisonal la predicción puede intervenir como un elemento configurador de la realidad futura, lo que no ocurre en el hecho de que mañana llueva o no llueva. Si predigo que mañana lloverá, mi predicción, como la creencia de las personas en mi predicción, no influye en el comportamiento atmosférico. Sin embargo, en el ámbito jurídico, las predicciones se comportan de una forma distinta. Las predicciones pueden ejercer una función performativa en la medida en que sean tomadas como reglas o criterios que han de ser tenidos en cuenta en la toma de una decisión que afecta a la persona o grupo de personas sobre cuya conducta se hace la predicción. A su vez, la confianza en las predicciones puede llevar a tomarlas como elemento de fundamentación de la decisión.

La automatización de decisiones basadas en predicciones, o fuertemente influidas por predicciones, puede llevar a que la autoridad decisonal tienda a conceder veracidad a la previsión establecida por el sistema predictivo. Esta posibilidad se puede ver reforzada por la generación de un estado de opinión en el que la credibilidad del sistema predictivo sea superior a la del juicio humano, o incluso que sea de observancia obligatoria en el caso de que el uso del sistema de IA fuera impuesto legalmente.

⁴⁷ TONRY, M., «Sentencing and Prediction: Old Wine in Old Bottles», en Keijser, J. W De; Roberts, J. V. y Jesper, R. (eds.), *Predicting Sentencing. Normative and Empirical Perspectives*, Oxford, Hart Publishing, 2020, pp. 269-298, p. 274.

Virginia Eubanks⁴⁸ ha analizado críticamente el uso que algunos servicios sociales en EE. UU. hacen de sistemas algorítmicos con funciones predictivas en la toma de decisiones sobre la concesión o denegación de ayudas a personas solicitantes. En su argumentación, Eubanks expone que, si el estudio de casos que desarrollan los servicios técnicos no se ajusta al resultado emitido por el algoritmo, la profesional piensa que algo se le ha pasado por alto. En caso de colisión, la profesional tiende a confiar en el criterio de la máquina⁴⁹ siempre que suponga que la herramienta algorítmica es más fiable que el cribado humano⁵⁰. En la práctica, «el algoritmo parece ejercer la función de instruir al personal de cribado»⁵¹. Lo que en apariencia se mostraría como una actividad científica basada en evidencias objetivas e infalibles es en realidad la traslación de la teorización que han elaborado un grupo reducido de profesionales: «Los datos de validación del modelo son un compendio de decisiones tomadas por asistentes sociales, investigadores y jueces, y llevan la impronta de sus prejuicios. [...] El criterio automatizado del modelo predictivo, en realidad, es el criterio de unas cuantas personas»⁵².

A la vista de estos argumentos, uno de los riesgos que puede introducir no solo el uso de los sistemas predictivos, sino también su automatización⁵³, es la limitación del ámbito y el material de decisión con el que trabajan los decisores. Los juicios predictivos pueden contener juicios valorativos previos no explícitos que quedan subsumidos en la predicción. Al actuar según la predicción, estaríamos validando esa base valorativa que queda oculta en la predicción. Por este motivo, es preciso analizar hasta qué punto la predicción se realiza en base a valoraciones que pueden quedar ocultas bajo la apariencia de objetividad del sistema predictivo.

2.3 Entre la «variabilidad no deseada» y la necesidad de variabilidad

Cuando hablamos de comportamientos humanos, no podemos saber por adelantado cuál será la conducta futura de una persona, ni

⁴⁸ EUBANKS, V., *La automatización de la desigualdad. Herramientas de tecnología avanzada para supervisar y castigar a los pobres*, Madrid, Capitán Swing, 2021.

⁴⁹ *Se trata del fenómeno conocido como «sesgo de automatización».* Vid. RIVAS, P., «Sesgos de automatización y discriminación algorítmica», en Rivas, P. (dir.), *Discriminación algorítmica en el ámbito laboral*, Cizur Menor, Aranzadi, 2022, pp. 31-68.

⁵⁰ EUBANKS, V., *op. cit.*, p. 171.

⁵¹ *Idem*, p. 172.

⁵² *Idem*, p. 200.

⁵³ Por «automatizar» se entiende la sustitución del operador humano por sistemas computacionales, pero al mismo tiempo «automatizar» expresa la idea según la cual el sistema opera por sí solo, sin intervención directa del operador, y, en tercer lugar, también significa que en un proceso decisonal dado parte de los contenidos que son tenidos en cuenta para tomar una decisión son volcados directamente del sistema al proceso decisonal.

cuáles serán las circunstancias que se darán. Hart planteaba la imposibilidad de una teoría jurídica «mecánica», ya que nos es imposible conocer la variabilidad de rasgos que intervienen en los hechos posibles futuros, así como las combinaciones que se pueden dar entre estos rasgos. Por esto, señalaba Hart, no es posible avanzar reglas que especifiquen la solución a los problemas futuros con los que nos podemos encontrar⁵⁴.

Si bien Hart defendió que lo razonable era atender las cuestiones futuras cuando «se presentan y son identificadas»⁵⁵, y de esta forma no confundir la voluntad de certeza y predictibilidad con la tendencia a «prejuizar ciegamente lo que ha de hacerse en un campo de casos futuros, cuya composición ignoramos»⁵⁶, lo cierto es que en ocasiones necesitamos adelantarnos a los hechos que están por venir. Ante la incapacidad humana para anticipar el futuro, los sistemas jurídicos han atendido esta incapacidad mediante una variedad de técnicas⁵⁷. Los sistemas predictivos forman parte de esta necesidad de anticiparse al futuro.

Kahneman *et al.* utilizan la categoría «variabilidad no deseada» para referirse a decisiones que se toman en contextos decisionales dispares como son el de las compañías aseguradoras, el de los tribunales penales o el de las diagnósticos médicos. Esta disparidad resta precisión al uso que estos autores hacen de esta expresión. Se puede estar de acuerdo en que hay situaciones en las que se puede aplicar con facilidad una regla o reglas conectadas entre sí y también se puede acordar que en tales situaciones es deseable hacerlo. Sería el caso de multar por exceso de velocidad. Sin embargo, la propuesta que hace Kahneman *et al.* en el sentido de utilizar los algoritmos para mejorar la decisión humana presenta problemas si se aplica a juicios en los cuales la decisión ha de ser tomada a partir de una evaluación de circunstancias y, en su caso, mediante la ponderación entre las argumentaciones aplicables en la resolución del caso.

La automatización de la decisión de multar por exceso de velocidad, así como la gestión de la multa impuesta, crea un escenario en el que, siempre que los radares utilizados estén correctamente calibrados y se observen los plazos de notificación (entre otras cuestiones objetivables), el procedimiento a seguir puede ser sometido con facilidad a reglas. En este tipo de casos, es relativamente fácil identificar cuál sería la variabilidad indeseable ya que disponemos de reglas aplicables y, por lo tanto, una decisión modelo que se ha de aplicar a cada caso que cumple las circunstancias estipuladas en los supuestos de hecho de las reglas. La decisión modelo, que se ajustaría idealmente al

⁵⁴ HART, H. L. A., *El concepto de Derecho*, Buenos Aires, Abeledo-Perrot, 1963, p. 160.

⁵⁵ HART, H. L. A., *op. cit.*, p. 162.

⁵⁶ *Ibidem.*

⁵⁷ HART, H. L. A., *op. cit.*, p. 163.

contenido normativo aplicable, se explica a partir de la combinación secuencial entre dos componentes de la regla aplicable: el «predicado fáctico» mediante el que se especifican las condiciones fácticas que dan lugar a la aplicación de la regla y la prescripción de lo que ha de ocurrir cuando se verifiquen las condiciones especificadas en el predicado fáctico, el «consecuente»⁵⁸.

Sin embargo, muchos otros ámbitos de decisión jurídicamente relevantes presentan una textura abierta en los que la utilización de sistemas predictivos presenta problemas. Pongamos el ejemplo de la decisión sobre la guarda y custodia de los hijos e hijas en un proceso de separación o divorcio. En 2021, hubo en España 86.851 divorcios, 3.674 separaciones y 57 nulidades. En el 53,1 % de casos se otorgó la custodia a la madre, en el 3,5 % al padre, y en el 43,1 % de los casos se decidió que la custodia fuera compartida. En el 0,3 %, la custodia fue otorgada a otras instituciones o a familiares⁵⁹.

Dado que en la toma de esta decisión concurren diferentes aspectos a tener en cuenta, supongamos que se da prioridad al bienestar del menor como principio a preservar en la toma de la decisión sobre la adjudicación de la guarda y custodia. La variabilidad no deseada se daría cuando se tomase una decisión de adjudicación de guarda y custodia, en base a sesgos o ruidos (siguiendo con la terminología utilizada por Kahnemann *et al.*) que perjudicara el bienestar del menor. Bajo esta premisa, se podría proponer la automatización de decisiones que evitara las variabilidades no deseadas en la toma de la decisión. Para desarrollar este sistema habría que elegir qué predictores hay que utilizar y qué peso hay que conceder a cada predictor. Para seleccionar los predictores y su valor, tanto individualmente considerado como en su interacción con otros predictores, habría que partir de una teorización acerca del bienestar infantil. En concreto, habría que establecer cuáles son los factores que contribuyen a su bienestar y cuáles son los factores de riesgo que hay que prevenir. De esta forma, el sistema podría emitir un resultado que predijera cómo se garantiza, en términos de probabilidad, el bienestar del menor. En este tipo de predicción mecanizada, se combinaría la predicción sobre un acontecimiento futuro con la decisión evaluativa acerca de cuáles son los factores que se considera que inciden en el bienestar del menor y cuál es el grado de cumplimiento que puede garantizar cada uno de los progenitores.

La *causalidad probabilística* no equivale a la certeza sobre un hecho futuro. En el caso de los sistemas automatizables utilizados para la formulación de juicios predictivos, lo que es una causalidad

⁵⁸ SCHAUER, F., *Las reglas en juego. Un examen filosófico de la toma de decisiones basadas en reglas en el derecho y en la vida cotidiana*, Barcelona, Marcial Pons, 2004, p. 81.

⁵⁹ Instituto Nacional de Estadística. Estadística de nulidades, separaciones y divorcios. Resultados. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176798&menu=resultados&idp=1254735573206

probabilística⁶⁰ puede presentarse, o ser entendido por quienes van a utilizar el sistema para fundamentar una decisión, no como una probabilidad sino como una relación causal entre los datos de los que se nutre el sistema y las consecuencias que han quedado vinculadas a la explotación de estos datos. De darse este fenómeno, se estaría ante la automatización, ya no solo del sistema predictivo sino también del valor atribuido a la predicción en el proceso decisional. Por esta deriva, se correría el riesgo de tratar una causalidad probabilística como una causalidad directa⁶¹. Como señala Schauer: «Decir que existe una causalidad probabilística entre una propiedad (perro) y una consecuencia (molestias) no es lo mismo que decir que la ocurrencia de la propiedad es necesariamente seguida por la ocurrencia de la consecuencia, ni tampoco es decir que el acaecimiento de la consecuencia requiere la ocurrencia de la propiedad. Simplemente es decir que la incidencia de la consecuencia será mayor en una población que posee la propiedad en cuestión que en una población idéntica en todos los demás aspectos en el cual esta propiedad está ausente»⁶².

3. EL ESTATUS NORMATIVO DE LOS SISTEMAS PREDICTIVOS UTILIZADOS POR LAS AUTORIDADES EN PROCESOS DECISIONALES

El uso de sistemas predictivos por parte de las autoridades lleva a preguntarse cuál es su estatus normativo. Desde el punto de vista del control de legalidad de la actuación de los poderes públicos, cabe entender que los sistemas utilizados han de ser validados previamente a su utilización y que la aplicación de sus resultados ha de ser susceptible de ser controlada legalmente⁶³. La validación legal de un sistema debería abarcar cuestiones como: la delimitación del ámbito decisio-

⁶⁰ BUNGE, M., *A la caza de la realidad. La controversia sobre el realismo*, Barcelona, Gedisa, 2007, p. 150 ss. Vid. tb. TARUFFO, M., en su análisis crítico sobre las limitaciones de los modelos probabilísticos aplicados a la prueba judicial, tanto de los modelos bayesianos (también criticados por Bunge) como los de probabilidad cuantitativa, *La prueba de los hechos*, Madrid, Trotta, 2002, pp. 192-193, 215-216 y, en la entrevista final a cargo de Jordi Ferrer, pp. 529-530.

⁶¹ SCHAUER, F., *op. cit.*, p. 88.

⁶² *Ibidem*.

⁶³ BOIX PALOP, A. parte de la premisa según la cual el código fuente integrado en la adopción de decisiones administrativas tiene materialmente valor normativo de reglamento y que «los algoritmos empleados por la Administración pública de modo no puramente instrumental producen materialmente los mismos efectos que cualquier reglamento, al preordenar la decisión final del poder público y limitar el ámbito de discreción o de capacidad de determinación de quienes los han de aplicar a partir de los postulados contenidos en la programación», «Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones», *Revista de Derecho Público: teoría y método*, vol. I, 2020, pp. 223-270, p. 237.

nal en el que se aplicará el sistema, qué autoridad lo puede utilizar, en el ejercicio de qué funciones, la vinculatoriedad de los resultados emitidos por el sistema, qué factores son tenidos en cuenta y con qué peso al predecir un hecho futuro o qué relación establece el sistema entre los distintos factores preseleccionados como relevantes para predecir un hecho futuro.

De otra forma, nos hallaríamos ante un cuerpo extraño al ordenamiento jurídico que sería utilizado sin título de validez legal. Sin embargo, puede suceder, por vía de hecho, que el uso de un sistema de IA, por ejemplo, para la predicción del riesgo de reincidencia, responda a la propia actuación de la autoridad que ha incorporado sistemas de gestión a los que se han ido añadiendo funciones conforme se desarrollaban nuevas tecnologías que ampliaban las funciones de estos sistemas.

Tomemos como ejemplo un sistema que se utiliza actualmente en el cálculo y la gestión del riesgo de maltrato a menores. La administración pública catalana tiene disponible en abierto un simulador para ayudar a gestionar el riesgo de maltrato infantil (MSGR por sus siglas en catalán⁶⁴). En las indicaciones de uso del sistema se señala que este sistema no está pensado «para sustituir la decisión que ha de tomar el/la profesional»⁶⁵. El engarce legal de este sistema, tanto en su diseño, como en su entrenamiento y utilización, se sustenta en una Orden por la que se autoriza la creación de instrumentos de apoyo tecnológico para facilitar la valoración general de la situación de riesgo a partir de los «indicadores y factores objetivados y, si fuera necesario, establecer la orientación de las acciones que se recomienda tomar»⁶⁶. Esta disposición legal establece qué tipo de información ha de ser tenida en cuenta para evaluar el riesgo de maltrato presente y futuro. Se establecen indicadores de maltrato (120 indicadores que quedan organizados en distintos bloques), factores de riesgo/protección (51 indicadores), situaciones de riesgo y desamparo (86 indicadores), situación personal, familiar y social del niño-adolescente (30 indicadores), aspecto físico y situación personal del niño o adolescente (67 indicadores), área emocional y comportamental del niño o adolescente (56 indicadores), desarrollo y aprendizaje (16 indicado-

⁶⁴ Módulo de apoyo a la gestión del riesgo (MSGR): https://drets-socials.gencat.cat/ca/ambits_tematics/infancia_i_adolescencia/proteccio_a_la_infancia_i_adolescencia/maltractaments_dinfants_i_adolescents/registre_unificat_de_maltractaments_infantils_rumi

⁶⁵ Puede consultarse en la web citada anteriormente. Ver igualmente los documentos disponibles en esta misma dirección digital: «Listado de observaciones» (documento que incluye los indicadores que toma en consideración el sistema algorítmico) y el «Manual de uso».

⁶⁶ Orden BSF/331/2013, de 18 de diciembre, por la que se aprueban las listas de indicadores y factores de protección de los niños y adolescentes (DOGC núm. 6530 de 30 de diciembre de 2013), que desarrolla diferentes previsiones establecidas en la Ley catalana 14/2010, de 27 de mayo, de los derechos y las oportunidades de la infancia y la adolescencia.

res), relaciones de los progenitores, tutores o guardadores con el niño o adolescente (114 indicadores), contexto social, familiar y personal del niño o adolescente (41 indicadores). A partir de estos indicadores, los profesionales evalúan en qué grado se dan estos indicadores en cada situación concreta. Una vez introducida esta información en el sistema, el sistema emite una valoración sobre la gravedad de la situación, la tipología del maltrato, el riesgo de repetición y las acciones que se recomienda tomar.

En el caso del sistema RisCanvi, que ha sido presentado más arriba, el análisis sobre la validez legal del sistema presenta más confusión que el sistema MSGR. La normativa en materia de concesión de permisos de salida penitenciaria establece, entre otras disposiciones, que se pueden conceder permisos de salida previo informe preceptivo del Equipo técnico. La normativa vigente establece qué características ha de tener este informe y bajo qué circunstancias ha de ser desfavorable⁶⁷: «la peculiar trayectoria delictiva, la personalidad anómala del interno o la existencia de variables cualitativas desfavorables» que hagan pensar que es probable el «quebrantamiento de la condena, la comisión de nuevos delitos o una repercusión negativa de la salida sobre el interno desde la perspectiva de su preparación para la vida en libertad o de su programa individualizado de tratamiento»⁶⁸.

Desde el punto de vista de la formación de un juicio sobre la probabilidad de que suceda un hecho futuro, la normativa citada presenta un gran interés, tanto como su dificultad de determinación. Pese a la dificultad que supone dotar de contenido preciso (tan preciso como el que se necesite para tomar una decisión) a expresiones como «peculiar trayectoria delictiva», «la personalidad anómala del interno» o «existencia de variables cualitativas desfavorables», lo cierto es que estos informes han de ser emitidos, y que los informes han de ser razonados. Es en relación con esta tarea que se han ido generando protocolos o sistemas de apoyo para orientar la toma de la decisión⁶⁹.

⁶⁷ Tiene igualmente especial interés recordar que el Equipo técnico puede estar formado por un jurista, un psicólogo, un pedagogo, un sociólogo, un médico, un ayudante técnico sanitario, un maestro o encargado de taller, un educador, un trabajador social, un monitor sociocultural o deportivo y un encargado de departamento (art. 274 del Real Decreto 190/1996, de 9 de febrero, por el que se aprueba el Reglamento penitenciario). A partir de esta configuración, cabe entender que el informe del Equipo Técnico se configura de forma colaborativa a partir de los juicios emitidos de forma individual por cada miembro del equipo y que son puestos en común para tomar una decisión colectiva.

⁶⁸ *Vid.* Artículo 156, del Real Decreto 190/1996, de 9 de febrero, por el que se aprueba el Reglamento penitenciario.

⁶⁹ Cabría preguntarse si estos sistemas contienen normas técnicas o si los informes elaborados a partir ellos son normas técnicas. Por el momento, no existen normas técnicas que regulen los requisitos técnicos que han de cumplir los sistemas predictivos, como sí que existen normas técnicas en materia de edificación o de industria, por ejemplo, *vid.* TARRÉS VIVES, M., *Normas técnicas y ordenamiento jurídico*, Valencia, Tirant lo Blanch, 2003, p. 24, p. 237 ss.; *vid.* tb. ESTEVE PARDO, J., *Técnica, riesgo y*

3.1 La prohibición de automatización de las decisiones individualizadas

Tanto en el derecho comunitario como en el derecho español, durante los últimos años se ha legislado en materia de la «automatización de decisiones». Por decisión automatizada se entiende: aquella decisión que se basa únicamente en el tratamiento automatizado de datos y que se toma mediante un instrumento tecnológico sin que haya participación humana en el proceso de decisión⁷⁰.

El Parlamento europeo y el Consejo reguló el tratamiento y la libre circulación de datos mediante el Reglamento 2016/679⁷¹. Este Reglamento dedica su artículo 22 a las «Decisiones individuales automatizadas, incluida la elaboración de perfiles». La normativa vigente reconoce a los administrados el derecho a «no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar» (art. 22.1⁷²).

Derecho, Barcelona, Ariel, 1999, pp. 39-40; ESTEVE PARDO, J., *El desconcierto del Leviatán. Política y derecho ante las incertidumbres de la ciencia*, Madrid, Marcial Pons, 2009.

⁷⁰ Vid. «Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a efectos del Reglamento 2016/679», adoptadas el 3 de octubre de 2017, revisadas y adoptadas el 6 de febrero de 2018, a cargo del Grupo de trabajo sobre protección de datos del artículo 29, WP25 Irev.01., p. 8. Por ejemplo, la reciente Ley 15/2022, de 12 de julio, integral para la igual de trato y la no discriminación. En su artículo 3.1. o) determina la «Inteligencia artificial y gestión masiva de datos, así como otras esferas de análoga significación», como ámbito objetivo de aplicación de las disposiciones contenidas en esta ley. Esta misma ley dedica su artículo 23 a «Inteligencia artificial y mecanismos de toma de decisión automatizados».

⁷¹ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos, artículos 13.2.f, 14.2.g, 22.1, 22.4 y 47.2.e, entre otros. Con anterioridad, mediante la Directiva 95/46/CE, del Parlamento europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que afecta al tratamiento de datos personales y a la libre circulación de datos, ya se habían establecido las bases mínimas de la regulación que fue desarrollada con amplitud por el citado Reglamento 2016/679.

Se ha de tener presente también la Directiva 2016/680 del Parlamento europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos, artículo 11: «Mecanismo de decisión individual automatizado».

Vid. tb. la Propuesta de Reglamento del Parlamento europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión, 21 de abril de 2021, COM(2021) 206 final.

⁷² No obstante, se pueden utilizar datos agregados con fines estadísticos, siempre que los datos personales o los resultados de trabajar con datos agregados no «se utilicen para respaldar medidas o decisiones relativas a personas físicas concretas», (Considerando 162 del Reglamento). En España, la Ley Orgánica 3/2018, de 5 de diciem-

La prohibición de automatizar decisiones ha sido establecida legalmente de tal forma que, desde mi punto de vista, se puede hablar de una autorización de la automatización de decisiones, aunque con algunas limitaciones. Si se tiene en cuenta que se permiten las automatizaciones que sean necesarias para la celebración y la ejecución de un contrato, aquellas en las que consienta explícitamente la persona interesada⁷³, las que sean necesarias por razones de interés público esencial y las autorizadas por el Derecho de la Unión o de los Estados miembros, se llega a entender que el actual marco regulatorio es más comprensible como una autorización con limitaciones que como una prohibición con excepciones.

Para entender qué papel están desempeñando o pueden desempeñar las automatizaciones de los sistemas predictivos en los procesos de toma de decisiones jurídicamente vinculantes, se pueden analizar los derechos atribuidos a las personas que se pueden ver afectadas por estas decisiones y a las obligaciones, relacionados con estos derechos, asignadas a las entidades que hacen uso de las automatizaciones. Entre los distintos derechos y obligaciones actualmente previstos en el derecho comunitario y en el derecho español⁷⁴ se presta ahora atención a dos disposiciones: el derecho a saber y la prohibición de utilizar determinados tipos de datos en la toma de decisiones automatizadas. Ambas disposiciones establecen el derecho que la persona tiene a conocer la fundamentación de la decisión que se toma sobre ella.

3.2 El derecho a saber y la prohibición de utilizar determinados tipos de datos en la automatización de decisiones

Uno de los problemas que presenta la automatización de decisiones es que las personas afectadas pueden ignorar que se les están aplicando decisiones que han sido tomadas mediante mecanismos de automatización. Como criterio general, la regulación vigente permite la utilización de sistemas de IA en la toma de decisiones siempre que

bre, de protección de datos personales y garantía de los derechos digitales, artículo 11.2, reconoce el derecho de la persona a saber que puede ser objeto de una decisión individual automatizada y el derecho a oponerse a la adopción de esa decisión individual automatizada que puede producir efectos jurídicos sobre esta persona o le puede afectar significativamente.

⁷³ Considerando 71 del Reglamento 2016/679. *Vid.* «Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a efectos del Reglamento 2016/679», adoptadas el 3 de octubre de 2017, revisadas y adoptadas el 6 de febrero de 2018, a cargo del Grupo de trabajo sobre protección de datos del artículo 29, WP251rev.01., p. 22 ss.

⁷⁴ Para una visión amplia del conjunto de derechos de las personas ante las decisiones automatizadas que les puedan afectar y las obligaciones de los responsables de datos, *vid.* RODRÍGUEZ AYUSO, J. F., *Garantía administrativa de los derechos del interesado en materia de protección de datos personales*, Barcelona, Bosch Editores, 2021.

las decisiones no estén basadas «únicamente» en el tratamiento automatizado de datos. El término clave es «únicamente». Si una autoridad toma una decisión basándose en el resultado del proceso de automatización de datos, y complementa o matiza suficientemente este resultado con otros elementos de decisión, o argumenta a partir del resultado emitido por uno o varios sistemas de IA, se entendería que ya no se está ante una decisión basada «únicamente» en el tratamiento automatizado de datos.

La intervención ha de tener contenido real y no ser una mera apariencia. Esta previsión queda complementada por lo que debería constituir una salvaguarda en el proceso decisonal: el derecho a la intervención humana en la configuración de la decisión⁷⁵. Esta intervención en la toma de la decisión se reserva a una persona autorizada legalmente para modificar la decisión⁷⁶. Y, por lo tanto, en coherencia con esta previsión y como se acaba de señalar, quien interviene ha de estar en disposición de explicitar cuál ha sido el contenido y efecto de su intervención. Es muy probable que, en un futuro inmediato, la determinación de cuándo la intervención de la autoridad es suficiente o insuficiente para satisfacer el estándar legal exigido vaya a ser una de las cuestiones más controvertidas.

El derecho a saber se ha construido como un derecho complejo que contiene prescripciones como: el derecho a ser informados acerca de la utilización de sistemas automatizados de tratamientos de datos, el derecho a recibir «información significativa»⁷⁷ sobre la lógica que se sigue en la elaboración de la decisión y sobre los criterios aplicados en la toma de decisión⁷⁸ y el derecho a ser informados de la importancia y las consecuencias que estas decisiones automatizadas pueden tener para las personas interesadas⁷⁹.

El acceso a la información significativa sobre la lógica seguida en la elaboración de una decisión concreta ocupa una posición estratégi-

⁷⁵ Vid. «Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a efectos del Reglamento 2016/679», adoptadas el 3 de octubre de 2017, revisadas y adoptadas el 6 de febrero de 2018, a cargo del Grupo de trabajo sobre protección de datos del artículo 29, WP251rev.01., p. 30. En este mismo sentido, ver el punto 4.2. y 6.9. del «Dictamen del Comité Económico y Social Europeo sobre “Inteligencia artificial”: anticipar su impacto en el trabajo para garantizar una transición justa», (2018/C 440/01), 19 de septiembre de 2018.

⁷⁶ Vid. «Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a efectos del Reglamento 2016/679», adoptadas el 3 de octubre de 2017, revisadas y adoptadas el 6 de febrero de 2018, a cargo del Grupo de trabajo sobre protección de datos del artículo 29, WP251rev.01., p. 23.

⁷⁷ El derecho a recibir «información significativa» no incluye el acceso al código base del sistema algorítmico utilizado. La legislación vigente, así como la jurisprudencia, ha asumido que esta posibilidad entraría en colisión con la protección de la propiedad intelectual mediante la que se protegen los programas informáticos.

⁷⁸ En este sentido, ver el punto 4.2. del «Dictamen del Comité Económico y Social Europeo sobre “Inteligencia artificial”: anticipar su impacto en el trabajo para garantizar una transición justa», (2018/C 440/01), 19 de septiembre de 2018.

⁷⁹ Artículo 13.2.f, 14.2.g, 15.1.h del Reglamento 2016/679.

ca. Se entiende que se respeta el derecho a tener esta «información significativa» cuando quien avala la decisión informa acerca del tipo de datos que han sido utilizados al tomar la decisión, explica por qué los datos utilizados son más relevantes que otros datos disponibles, y explica, si es el caso, por qué y cómo se han generado y se han utilizado estas perfilaciones en la toma de la decisión y por qué es necesario utilizar ese perfil⁸⁰. Habría que añadir a esta lista de requisitos otra exigencia si se quiere garantizar el derecho de la persona a conocer la fundamentación de la decisión que le afecta directamente: la explicación de las conexiones que el sistema de IA realiza entre los distintos datos utilizados para formular una decisión. Estas conexiones forman parte del «razonamiento» que sigue o crea el algoritmo para tomar una decisión. «En ocasiones, incluso para quienes diseñan y entrenan un sistema avanzado de IA es difícil explicar por qué el sistema está funcionando como lo está haciendo»⁸¹.

4. A MODO DE CONCLUSIÓN: LA RESPONSABILIDAD POR LA CONFIGURACIÓN DE LA DECISIÓN

La incorporación de sistemas predictivos automatizables en los procesos de toma de decisiones supone una novedad en la arquitectura decisional utilizada por las autoridades públicas. Los sistemas predictivos, tanto por su funcionalidad pretendida, como por las funciones realmente alcanzadas, ofrecen modelos predictivos que orientan la toma de decisión en contextos de incertidumbre. Para las autoridades que han de tomar decisiones, estos sistemas proveen resultados utilizables en la configuración la decisión.

Dada la influencia que estos sistemas pueden tener en la actividad decisional, es preciso establecer una serie de salvaguardas que se muestran urgentes dado el desequilibrio existente entre el desarrollo y aplicación de la IA y los niveles de eficacia regulatoria existentes⁸². Estas salvaguardas pivotan, desde una perspectiva de control de legalidad, sobre el estatus normativo de los sistemas predictivos. Hasta el momento, estos sistemas aparecen como protocolos de evaluación y

⁸⁰ Vid. «Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a efectos del Reglamento 2016/679», adoptadas el 3 de octubre de 2017, revisadas y adoptadas el 6 de febrero de 2018, a cargo del Grupo de trabajo sobre protección de datos del artículo 29, WP251rev.01., p. 35.

⁸¹ Vid. PASQUALE, F., *The Black Box Society. The Secret Algorithms that Control Money and Information*, Cambridge, Harvard University Press, 2015, en especial p. 216 ss.

⁸² Como ejemplo, vid. la carta publicada por el *Future of Life Institute*, en marzo de 2023, en la que centenares de especialistas y aplicadores de IA piden una moratoria de por lo menos 6 meses en el desarrollo y aplicación de sistemas de IA: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

gestión. Sin embargo, la posición que se ha defendido en este texto es la siguiente: se hace necesario revisar el estatus normativo de estos sistemas dada la posición estratégica que ocupan y, previsiblemente, van a ocupar en un futuro inmediato. En concreto, y como mínimo, es necesario validar la base fáctica y evaluativa de los sistemas utilizados, especialmente cuando contienen valoraciones que van a quedar insertas en las predicciones, se ha de explicitar en términos comprensibles para la ciudadanía la racionalidad del sistema, es preciso asegurar la trazabilidad de la predicción que se utiliza para fundamentar una decisión, y, cuestión política y ética fundamental, hay que decidir qué enfoque se le quiere dar a los falsos positivos.

No se puede olvidar que estos requisitos legales pueden ser contemplados como una extensión adaptativa a la nueva realidad tecnológica de requisitos que ya estaban incorporados en nuestra tradición legal. Es el caso de la necesidad de la motivación de la decisión o la regulación legal de los medios de prueba, por ejemplo. Sin embargo, como se ha sostenido en este texto, la posibilidad de aplicar IA a los sistemas predictivos, lejos de excepcionar los requisitos legales ya presentes en nuestros ordenamientos jurídicos, hace necesaria su adaptación a la nueva realidad en la que nos hallamos inmersos.

Parte de la realidad en la que estos sistemas predictivos se van a utilizar durante los próximos años va a venir condicionada por la percepción mediática y social acerca de las consecuencias de las decisiones que se tomen siguiendo sistemas predictivos, pero también acerca de las consecuencias de las decisiones que se tomen desoyendo los resultados emitidos por estos sistemas. Ante este escenario en el que se debate acerca de la conveniencia, u obligación si así se estableciera legalmente, de utilizar predicciones automatizadas para tomar decisiones, surge con fuerza ya no solo la responsabilidad por la decisión tomada, sino también la responsabilidad por la configuración de la decisión. Y, en concreto, la responsabilidad por los sistemas predictivos automatizados insertos en el proceso decisional. Si se acepta como pieza fundamental de nuestro sistema jurídico que la responsabilidad por la decisión recae en la autoridad legalmente competente para tomar la decisión, esta responsabilidad no debería ser transferida en el sistema artificial. Sin embargo, esta formulación no agota el problema, especialmente cuando los sistemas, como se ha visto en este texto, pasan a ocupar una posición estratégica en la configuración y fundamentación de la decisión.