

ÍNDICE

1. El concepto de Inteligencia Artificial 1
2. Inteligencia artificial, entidad jurídica y responsabilidad..... 3
3. El uso de la Inteligencia Artificial predictiva en el sistema penitenciario..... 5

Desafíos de la Inteligencia Artificial en el ámbito penitenciario: Derechos Fundamentales y Justicia Penal

Laura María Cabello Pérez y Laura López Pascual

El presente escrito tiene por finalidad llevar a cabo un análisis de los desafíos relacionados con el uso de sistemas de Inteligencia Artificial (en adelante, IA) en el ámbito del derecho penitenciario. La especial relevancia de los derechos con los que trata este ámbito de la justicia hace que el uso de sistemas de IA deba tener una particular supervisión y unos requisitos de transparencia que aseguren que su impacto en la población penitenciaria y sus derechos no es negativo; por otro lado, su uso debe ser enmarcado dentro de la finalidad reinsertiva en la que se fundamente las instituciones penitenciarias.

1. El concepto de Inteligencia Artificial

Para comenzar analizando esta problemática es oportuno definir qué es la inteligencia artificial. Según la RAE esta es una “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”¹

Tradicionalmente estos sistemas informáticos han sido clasificados por la ciencia en dos categorías: “IA débil” e “IA fuerte”. La primera corresponde a “la ciencia e ingeniería que permite diseñar y programar ordenadores de forma que realicen tareas que requieran inteligencia”.² Inteligencia en este sentido hace referencia a una “especializada”, solo diseñada para un tipo de tarea (o múltiples pero limitadas, ejemplo clásico: jugar al ajedrez, pero no a las damas.) Su objetivo es ayudar al ser humano con actividades mentales en lugar de duplicarlas.

Por otra parte, la denominada IA fuerte es “aquella que permite replicar la inteligencia humana mediante máquinas”³. El desarrollo de la primera ha sido tan extenso estos últimos años que se tiende a pensar que aquello ante lo que nos encontramos cuando nos encontramos por ejemplo ante un robot humanoide⁴ es precisamente un modelo de IA fuerte. Pero a pesar la sofisticación de algunos sistemas, estos aún no cumplen el requisito

¹RAE

²Ramón Lopez de Mantaras y Pedro Meseguir Gonzalez, *¿Que sabemos de? Inteligencia Artificial,,* CSIC, Catarata, 2017, p.8

³Idem, p.9

⁴Business times, Meta plans major investment into AI-powered humanoid 2025, <https://www.businesstimes.com.sg/companies-markets/telcos-media-tech/meta-plans-major-investment-ai-powered-humanoid-robots>

de autonomía-autoconciencia (a su vez relacionados con otras características como la creatividad, el sentido común y la empatía).

Los sistemas de IA actualmente se organizan y evolucionan a través de diversos métodos: razonamiento basado en restricciones, redes bayesianas, planificación, aprendizaje automático, etc.⁵ Al mismo tiempo en la actualidad podemos diferenciar los sistemas entre aquellos que aprenden a través de métodos de *machine learning* (sistema formado por modelos matemáticos y que funciona a través de algoritmos) y *deep learning* (sistema basado en datos y algoritmos que utiliza cálculos creados en una estructura similar a lo que sería un sistema de redes neuronales debido a su profundidad,). Los algoritmos en los que se basan este tipo de aprendizaje “aplican en cascada un conjunto de transformaciones no lineales, de forma que cada nivel recibe en entrada la salida del nivel anterior y lleva a cabo una abstracción aprendiendo así representaciones cada vez más complejas. Por ejemplo, para aprender a reconocer sillas basándose en imágenes de muchas sillas, como datos de entrenamiento, el primer nivel recibirá como entrada los pixeles de una imagen y lo que aprende son características como, por ejemplo, líneas rectas con distintas orientaciones”⁶ y así hasta llegar al último nivel, donde a partir de contornos y características abstractas se llega a la silla completa. Debido a su complejidad, podemos presumir que se acercaría más a lo que podría ser una inteligencia humana. Además, la clave del *deep learning* es que “estas características cada vez más abstractas no están diseñadas manualmente por el programador del sistema, sino que son automáticamente extraídas mediante aprendizaje”⁷. Sin embargo, a pesar de su autonomía sigue sin ser una inteligencia humana, porque al mismo tiempo para poder hacer esta comparación, deberíamos ser capaces de definir qué es esta, concepto respecto al cual sigue sin haber una completa unanimidad social y científica.

En concreto, los sistemas de IA predictivos que se usan en los sistemas judiciales y penitenciarios se organizan a través de algoritmos de aprendizaje y estadísticos y son capaces de realizar análisis y predecir sobre posibles resultados y patrones. Su sistema se basa en el *machine learning*⁸. Para ello necesita una gran cantidad de datos, lo que significa que una muestra pequeña o sesgada puede condicionar el resultado de la predicción. Por esa razón, la calidad de los datos y su pertinencia debe ser analizada con rigurosidad. Es fundamental al mismo tiempo tener en cuenta que, como toda estadística o predicción, las probabilidades de que estas ocurran nunca son absolutas.

Sin embargo, como ventaja reseñable y debido a la rapidez con la que analiza los datos, y la eficacia en el rastreo, esos sistemas predictivos pueden reducir el tiempo que se invierte en este tipo de tareas mecánicas y de cálculo, y proporcionar un espacio a los

⁵Ramón Lopez de Mantaras y Pedro Meseguir Gonzalez, *¿Que sabemos de? Inteligencia Artificial*,, CSIC, Catarata, 2017, p.97

⁶Ramón Lopez de Mantaras y Pedro Meseguir Gonzalez, *¿Que sabemos de? Inteligencia Artificial*,, CSIC, Catarata, 2017, p104

⁷Idem

⁸Cloudflare, 2025 <https://www.cloudflare.com/es-es/learning/ai/what-is-predictive-ai/>

seres humanos para encargarse de dar sentido a los datos que de las predicciones se desprenden, tal y como se hacen en las ciencias sociales con las estadísticas.

2. Inteligencia artificial, entidad jurídica y responsabilidad

Actualmente la IA carece de las características necesarias para ser considerada como un ente similar al ser humano. Así mismo, nuestra legislación no considera a los sistemas de IA como sujetos de derecho. Esto significa que en caso de daño la responsabilidad corresponde a los creadores y empresas distribuidoras del producto. Es por ello por lo que en el contexto de la Unión Europea se ha desarrollado una Directiva relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial. La misma tiene como objetivo aportar seguridad jurídica a los perjudicados, identificando a los sujetos responsables y facilitando la carga de la prueba al denunciante, que se enfrenta a una doble complejidad como causa de la oscura estructura de la *black box* en tanto que esta dificulta la información acerca de “la culpa del agente que circula un sistema inteligente, su funcionamiento deficiente, el nexo causal”⁹, etc. Actualmente la propuesta de Directiva ha sido retirada por la Comisión¹⁰

Al mismo tiempo el ya aprobado Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial en junio de 2024 (Reglamento IA) establece los requisitos y obligaciones que tienen los diferentes sujetos en la cadena de valor. El reglamento se alinea con la perspectiva de inteligencia artificial como herramienta al servicio de un ser humano. En última instancia, es este quien tiene la tarea de hacerse cargo de su funcionamiento. Así se subraya la necesidad de que la toma de decisiones y el uso de estos sistemas recaiga en una elección humana, reflexiva y justificada que garanticen que los sistemas de IA sean seguros, éticos y fiables. En caso de un uso irracional y de una falta de una posterior supervisión por parte de la persona física o jurídica, estaríamos expuestos a distintos peligros relativos a la arbitrariedad de las conclusiones alcanzadas por medio del uso de sistemas algorítmicos. Por esa razón la normativa establece una serie de obligaciones de supervisión y transparencia vinculados al distinto grado de riesgo. Para sistemas de IA con riesgo mínimo o nulo no se prevé ninguna medida, para aquellas con riesgo limitado como los *chatbots*, se impone un requisito de transparencia (indicación de las fuentes de contenido). En cuanto a IAs con riesgo alto, como la conducción autónoma o la identificación biométrica se debe cumplir con unas pruebas de comprobación rigurosas, y una mayor transparencia y supervisión humana. Por último, el reglamento prohíbe los usos que puedan perjudicar la seguridad o los derechos de las personas como el reconocimiento de emociones en el lugar de trabajo y centros educativos, los sistemas de puntuación ciudadana, etc.

⁹Manuel Ortiz Fernández, *La «adaptación» del derecho de daños a la inteligencia artificial: la propuesta de Directiva sobre responsabilidad*, Universidad Miguel Hernández de Elche, 2024, p. 4

¹⁰<https://euractiv.es/section/tech/news/commission-plans-to-withdraw-ai-liability-directive-draw-mixed-reactions/>

Por todo ello cabe apuntar a que, la inteligencia artificial a día de hoy es una herramienta, puesta al servicio del ser humano, el cual debe ser consciente de las ventajas y limitaciones que presentan y hacer un uso adecuado de las mismas.

No obstante, una de las dificultades es la denominada “black box” que mencionamos anteriormente y que se presenta especialmente en los sistemas de inteligencia artificial que utilizan modelos de *Deep learning* (DL). Estos sistemas son los que más dudas plantean sobre su fiabilidad.

El *Deep learning* se refiere a los modelos de aprendizaje artificial basados en sistemas algorítmicos, que, imitando los procesos neuronales biológicos, generan redes neuronales artificiales, a partir de las cuales se da el proceso de toma de decisiones. A pesar de la semejanza con los procesos neuronales humanos, paradójicamente se da una incomprensibilidad del referido proceso, ya que se caracteriza por ser no-lineal y opaco. Ello se traduce en que la conclusión a la que se llega en la toma de decisiones viene precedida por toda una compleja arquitectura algorítmica, cuyos patrones de razonamiento internos son los que, precisamente, nos resultan incomprensibles como humanos.

Por este motivo han ido surgiendo iniciativas para abrir la *black box* mediante el desgranamiento de los mencionados patrones, con el objetivo de hacer comprensible la arquitectura o red neuronal a partir de la cual una IA puede alcanzar determinada conclusión y no otra. Ahora bien, el problema que subyace a este tipo de iniciativas no es tanto la comprensión en sí de los patrones de razonamiento, cuanto la posibilidad de determinar si el uso de estos sistemas es confiable, si estamos a salvo de la arbitrariedad y, en su caso, cuáles son los problemas éticos que puedenemerger con el uso de modelos de razonamiento artificial.

El cuestionamiento de la relación de confianza que podemos depositar en los sistemas algorítmicos – relación denominada *e-Trust* – puede, por lo tanto, concretarse preguntándonos lo siguiente: ¿Es posible confiar en los sistemas algorítmicos que usan procesos de *Deep learning*? Y lo más relevante, ¿hasta qué punto podemos confiar en las consecuencias alcanzadas mediante un proceso de toma de decisiones algorítmico?

“El concepto estándar de confianza, según el cual una persona es considerada fiable y competente, no es suficiente para la confianza «moral». Necesitamos el requisito adicional de que las acciones se realicen de la manera correcta (cómo) y por las razones correctas (por qué)”¹¹. Así, podemos concluir que uno de los elementos fundamentales de la confianza es la transparencia, tanto en el proceso como en los motivos que fundamentan la decisión. Sin embargo, la opacidad y consecuente incomprensibilidad no sólo se refiere al proceso de razonamiento y de toma de decisiones. Por lo tanto, otro de los factores a considerar es el riesgo en el sesgo cognitivo, adquirido a través del *entrenamiento* de la red neuronal artificial.

¹¹Von Eschenbach, W. J., “Transparency and the Black Box Problem: why we do not trust AI”, *Philosophy and Technology*, (2021) p. 191

No podemos olvidar cómo este último aspecto tiene una especial repercusión en el uso de la IA para predecir el riesgo en ámbitos de la justicia penal, como el sistema penitenciario. A este respecto, no podemos obviar cómo “un análisis reciente [...] determinó que los acusados negros tenían muchas más probabilidades de ser juzgados incorrectamente como de alto riesgo de reincidencia que los acusados blancos, y los acusados blancos tenían muchas más probabilidades de ser juzgados incorrectamente como de bajo riesgo de reincidencia que sus homólogos negros. El análisis también mostró que «incluso cuando se controlaban los delitos anteriores, la reincidencia futura, la edad y el sexo, los acusados negros tenían un 45% más de probabilidades de que se les asignara una puntuación de riesgo más alta que a los acusados blancos»”¹².

3. El uso de la Inteligencia Artificial predictiva en el sistema penitenciario

La característica opacidad del proceso de toma de decisión de un sistema de IA tiene una especial relevancia en procedimientos penitenciarios como los permisos ordinarios, en los que existe un carácter discrecional en la decisión, una vez cumplido los requisitos objetivos (art. 154 y ss. del Reglamento Penitenciario)

El permiso ordinario penitenciario se justifica en base a la concepción del permiso como “mecanismo que promueve la reintegración social, al permitir mantener los vínculos familiares y sociales con el exterior; en que contribuyen a humanizar las penas, al aceptarse que la privación total de libertad sin ninguna esperanza de salir al exterior constituye un castigo inhumano; y en su utilidad para mantener la convivencia y el orden dentro de prisión.”¹³

Los permisos ordinarios (en contraposición con los extraordinarios por fallecimiento, alumbramiento, etc.) vienen regulados en el art. 47.2 de la Ley Orgánica General Penitenciaria y tienen como objetivo la preparación para la vida en libertad. Estos permisos se conceden en base a dos requisitos, los de tipo objetivo, a saber, que el preso esté en un segundo o tercer grado, y que se haya extinguido la cuarta parte de la condena. Por otro lado, es necesario otro requisito de carácter subjetivo, que no se observe mala conducta del preso, para lo que será necesario un informe previo del equipo técnico.

La subjetividad de este requisito reside en la valoración que se hace de las características y comportamiento del preso. Así, tal y como dispone el art. 156 del Reglamento Penitenciario “el informe del equipo técnico será desfavorable cuando por la peculiar trayectoria delictiva, la personalidad anómala del interno o por la existencia de variables cualitativas desfavorables, resulte 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

¹²Von Eschenbach, W. J., “Transparency and the Black Box Problem: why we do not trust AI”, *Philosophy and Technology*, (2021) p. 1912.

¹³Elena Larrauri, *¿Qué criterios se deben valorar en la concesión de permisos de salida?*, Jueces para la democracia, 2019, p. 43

de tratamiento.” De la calificación de estos extremos se encargan psicólogos, juristas, educadores sociales, etc. Y será posteriormente catalogado como favorable por la Junta de Tratamiento. En última instancia, viene otorgado por el Centro Directivo o el Juez de Vigilancia Penitenciaria (dependiendo de si el permiso supera las 48 horas). El programa RiscCanvi, ejemplo de uso de sistema de inteligencia artificial predictivo en el ámbito penitenciario, se utiliza de forma complementaria, “lo cual no quiere decir, que en ocasiones el sistema judicial revierta la decisión de la Junta alegando precisamente que la valoración del riesgo es alto”.¹⁴

En sistemas judiciales como el nuestro, colapsado por un incremento constante de los asuntos a abordar¹⁵, el uso de la IA puede suponer un apoyo aportando una mayor eficiencia a la hora de proporcionar datos, una reducción de los tiempos administrativos, automatización de procesos, reducción de costes, etc.

Estas funciones pueden ser de gran ayuda para evitar situaciones indeseadas en el entorno penitenciario, ya que “cuando los sistemas penitenciarios están desbordados y mal gestionados, las prisiones corren el riesgo de transformarse en lugares peligrosos para las personas detenidas y para el personal penitenciario, e incluso pueden convertirse en “escuelas del delito” y en ámbitos propicios para la radicalización”¹⁶ Sin embargo su uso debe ser siempre supervisado y revisado en última instancia por un ser humano.

Los sistemas de IA que ayudan a determinar el porcentaje de riesgo son sistemas predictivos de cálculo de probabilidad que tienen en cuenta distintos factores de riesgos como el comportamiento en el centro, el tipo de delito, su contexto socioeconómico, etc. Los datos son aportados por los profesionales antes mencionados. “Del sistema se extrae un resultado, que en ocasiones puede sugerir lo contrario a lo recomendado por los psicólogos o educadores sociales. Sin embargo, el juez y el fiscal, para la toma final de su decisión, únicamente reciben la puntuación final, lo que les impide valorar de forma individualizada a cada uno de los presos, y sus circunstancias.”¹⁷ Esta información, como ya se mencionó, es complementaria al resto de opiniones arrojadas en el informe. A pesar de ello, en ocasiones un porcentaje que arroje como resultado un alto riesgo puede significar un fuerte sesgo para la denegación de un permiso por parte del Juez o Centro Directivo. La dificultad entonces reside en poder acceder a un seguimiento de los datos y la conclusión a la que ha llegado el sistema de IA para arrojar ese porcentaje y poder así justificar la categorización del alto riesgo. La *black box* inherente a estos sistemas impiden el rastreo de la toma de decisiones, más que si la evaluación de riesgo fuese elaborada por un ser humano. Así mismo la decisión basada en estadísticas reemplaza al

¹⁴Elena Larrauri, *¿Qué criterios se deben valorar en la concesión de permisos de salida?*, Jueces para la democracia, 2019, p. 50

¹⁵<https://noticias.juridicas.com/actualidad/noticias/19873-casi-cuatro-millones-de-asuntos-pendientes-agravan-el-colapso-judicial-pese-a-que-se-dictan-un-1-1-mas-de-sentencias/>

¹⁶Oficina de las Naciones Unidas contra la Droga y el Delito, Abordando la crisis penitenciaria a nivel global, estrategia 2015-2017 https://www.unodc.org/documents/justice-and-prison-reform/Prison_Crisis_Strategy_Brochure_Spanish.pdf

¹⁷<https://www.fieldfisher.com/es-es/locations/espana/actualidad/algoritmo-para-decidir-sobre-la-reincidencia-de-los-presos-riscanvi>

caso concreto, no pudiendo expresar con adecuación la realidad compleja de la situación y comportamiento de una persona en prisión.

Por otro lado y con la publicación del Reglamento IA este tipo de sistemas predictivos deberán cumplir con una serie de obligaciones específicas en tanto el art. 6 y anexo III de la norma dispone que los “sistemas de IA destinados a ser utilizados por las autoridades garantes del cumplimiento del Derecho, o en su nombre, o por las instituciones, órganos y organismos de la Unión, en apoyo de las autoridades garantes del cumplimiento del Derecho para evaluar el riesgo de que una persona física cometa un delito o reincida en la comisión de un delito atendiendo no solo a la elaboración de perfiles de personas físicas mencionada en el artículo 3, punto 4, de la Directiva (UE) 2016/680 o para evaluar rasgos y características de la personalidad o comportamientos delictivos pasados de personas físicas o colectivos” son sistemas de alto riesgo.

El art. 26 del Reglamento establece que de utilizar estos sistemas se deberá “asignar supervisión humana, garantizar que los datos de entrada sean pertinentes y supervisar el funcionamiento del sistema. Si se detecta un riesgo, debe informarse al proveedor y a las autoridades competentes. Además, deben conservar los registros generados por el sistema de IA durante al menos seis meses.”¹⁸ Por otro lado el artículo 26 también establece que “los implantadores de sistemas de IA de alto riesgo que sean autoridades públicas o instituciones, órganos u organismos de la Unión cumplirán las obligaciones de registro a que se refiere el artículo 49. Cuando dichos implantadores descubran que el sistema de IA de alto riesgo que tienen previsto utilizar no ha sido registrado en la base de datos de la UE a que se refiere el artículo 71, no utilizarán dicho sistema e informarán de ello al proveedor o al distribuidor”.

También en octubre de 2024, el Consejo de Europa publicó diversas recomendaciones a sus Estados Miembros para que en el ámbito de los servicios penitenciarios se respeten los derechos fundamentales cuando se utilicen sistemas de inteligencia artificial. Su uso, señala el Consejo, debe ser legítimo y proporcionado. Sólo debe utilizarse, subraya, si estos sistemas contribuyen a la rehabilitación de los presos. Esta última finalidad del sistema de prisión no es ajena al ordenamiento jurídico español. Nuestra Constitución de 1978 establece en su art. 25 que “las penas privativas de libertad y las medidas de seguridad estarán orientadas hacia la reeducación y reinserción social”.

Además, esta reinserción social contiene dos exigencias, “a saber, que las penas no sean de excesiva duración, de modo que no se pueda transformar la finalidad constitucional en ilusoria o incluso que anique a la persona; y, por otra parte, que se fomente en todo caso el contacto del individuo con la sociedad en la que deberá integrarse algún día.”¹⁹ La primera exigencia viene salvaguardada por el uso residual de la prisión permanente (en el caso español, *revisable*). La segunda, por la inclusión en nuestro ordenamiento jurídico

¹⁸<https://artificialintelligenceact.eu/es/article/26/>

¹⁹Daniel Fernández Bermejo, *El fin constitucional de la reeducación y reinserción social ¿un derecho fundamental o una orientación política hacia el legislador español?* UDIMA, 2014, p. 378

de los permisos ordinarios, los cuales tienen como finalidad esa toma en contacto con la sociedad a la que el preso está destinado a reincorporarse.

Sin embargo, no solo basta la recopilación de un principio en un texto legal para que este sea llevado a cabo y respetado. Se necesita, asimismo, utilizar las herramientas materiales necesarias y una traducción en la realidad de estos principios. En este sentido, los sistemas predictivos de inteligencia artificial deben ser utilizados con el propósito de la reinserción, y no como un porcentaje numérico que impida al preso disfrutar de un derecho esencial dentro de nuestro sistema jurídico. Si las herramientas no sirven a sus propósitos las mismas deben ser analizadas y revisadas para que puedan tener un papel de optimización dentro del sistema.

No se niega el uso de estos sistemas, que pueden ser un apoyo para sistemas judiciales y penitenciarios, sino que se pone de relieve la necesidad de una mayor diligencia en su uso. Así la Recomendación del Consejo de Europa “señala que la IA puede ayudar a mantener la seguridad en las prisiones y, por lo tanto, permitir una mejor gestión de los riesgos y las crisis. Sin embargo, su uso debe ser estrictamente necesario y evitar efectos adversos en la intimidad y el bienestar de los delincuentes y del personal.”²⁰

Si los sistemas se utilizan, como es en el caso de España, para tomar decisiones acerca de la concesión o denegación de permisos, se deben asegurar medidas que eviten los sesgos en la toma de decisiones para evitar discriminación contra individuos o grupos de personas.

Un ejemplo de cómo funcionan estos sesgos es el siguiente²¹: cuando los sistemas predictivos se utilizan en actuaciones policiales, debido a que históricamente a ciertos grupos, perfiles, áreas, se les realizan de forma más habitual registros policiales, esto provoca que el algoritmo de predicción pronostique que ocurrirán más delitos en el futuro, lo que al mismo tiempo da como resultado una mayor presencia policial en estas zonas que en otras.

Y es que “existe una noción tan persistente como perjudicial de que la tecnología es neutra y objetiva: La percepción pública suele ser que la tecnología es intrínsecamente neutra y objetiva; algunos han señalado que esta presunción de objetividad y neutralidad tecnológica es una característica que aún sigue siendo una de las más destacables incluso entre los creadores de tecnologías. Ahora bien, la tecnología nunca es neutra, pues refleja los valores e intereses de quienes influyen en su diseño y uso, y está fundamentalmente conformada por las mismas estructuras de desigualdad que operan en la sociedad”²².

Este tipo de sesgos podría también aplicarse al uso de sistemas predictivos en la concesión de permisos. Características como el entorno social, la relación con la familia, las

²⁰Recommendation CM/Rec(2024)5 of the Committee of Ministers to member States regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services – Explanatory Memorandum

²¹<https://www.ohchr.org/es/stories/2024/07/racism-and-ai-bias-past-leads-bias-future>

²²Ashwini K. P., *Formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia*, 2024, p. 3

condiciones económicas, etc., vienen analizadas por el sistema como elementos que pueden influenciar en la reincidencia delictiva. Si bien es cierto que la criminología sostiene que los factores sociales y económicos son factores determinantes en la comisión de delitos, utilizar estos factores para juzgar la peligrosidad futura de un individuo, supone una suerte de tautología o circularidad que no refleja la verdadera naturaleza de la persona y las posibilidades de reinserción social de la misma. El sistema penitenciario español cree en la reinserción como objetivo último y por lo tanto lo valora como posible. Esto implica pensar que el preso está sujeto a cambio, es un ser humano que puede transformarse y mejorar, y que precisamente su estancia en prisión, la realización de actividades, y la concesión de permisos son elementos que pueden ayudar a este cambio. Por eso mismo “nuestro modelo penitenciario pretende vincular estos cambios de un modo muy concreto: la idea es que la persona evolucione según el recorrido tratamental propuesto —lo que implica determinados cambios— y, en caso de constatarse dicha evolución, es posible plantear la modificación de las condiciones de la ejecución para el disfrute de un mayor grado de libertad.”²³

El Consejo de Europa, en su recomendación y teniendo en cuenta los riesgos en el uso de estos sistemas, establece que las decisiones basadas en los sistemas de IA que puedan afectar a los derechos humanos deben estar sujetas a revisión humana y a mecanismos de denuncia. Por último, se establece el cuidado y la preservación de otro de los principios en los que se basa la reinserción, este es el del contacto humano como motor de cambio. Así “La IA y las tecnologías digitales relacionadas pueden ayudar a los profesionales a reintegrar a los delincuentes en la sociedad, pero las relaciones humanas deben seguir siendo un elemento central de la rehabilitación. En consecuencia, la IA y las tecnologías afines deben complementar y no sustituir las interacciones cara a cara -con profesionales de la salud, abogados, trabajadores sociales, otros profesionales y familias-, todas ellas esenciales para una rehabilitación eficaz.”²⁴

De no aplicar esta supervisión humanizada, podríamos enfrentarnos a una futura deshumanización de la justicia, de los procesos penitenciarios y al riesgo de caer en una suerte de determinismo social, el cual toma las circunstancias sociales de un sujeto como elemento determinante de su comportamiento en el futuro.

²³Lorena Alemán Aróstegui, *El Uso Del Riscanvi En La Toma De Decisiones Penitenciarias*, Departamento de Derecho. Facultad de Ciencias Jurídicas. Universidad Pública de Navarra, 2024, p.2

²⁴Recommendation CM/Rec(2024)5 of the Committee of Ministers to member States regarding the ethical and organisational aspects of the use of artificial intelligence and related digital technologies by prison and probation services – Explanatory Memorandum

