



El Estado algorítmico y la abdicación del criterio humano: sesgo de automatización y garantías constitucionales en los sistemas VioGén y RisCanvi

(The algorithmic state and the abdication of human judgement: Automation bias and constitutional guarantees in the VioGén and RisCanvi Systems)

OÑATI SOCIO-LEGAL SERIES VOLUME 16, ISSUE 5 (2026), 2234-2251: RECKONING WITH THE EAST: SITUATING SCHOLARSHIP FROM EASTERN EUROPE AND EURASIA IN CRIMINOLOGY AND SOCIO-LEGAL STUDIES

DOI LINK: <https://doi.org/10.35295/OSLS.IISL.2791>

RECEIVED 15 JUNE 2026, ACCEPTED 7 JULY 2026, FIRST-ONLINE PUBLISHED 20 JULY 2026, VERSION OF RECORD PUBLISHED 1 OCTOBER 2026

DANIEL MARTÍNEZ CRISTÓBAL¹ 

Resumen

Este artículo propone analizar, mediante una perspectiva sociojurídica interdisciplinar, el impacto que tienen los sistemas algorítmicos de evaluación de riesgos (VioGén y RisCanvi) en la Administración pública y en el sistema judicial en España. Se estudia cómo el uso de estas herramientas predictivas ha provocado que los profesionales dejen de lado su propio criterio debido al “sesgo de automatización” y por la presión que generan actualmente las dinámicas de protección y cautela en el seno de la Administración. Aplicando un enfoque cualitativo que integra fuentes normativas, evaluaciones técnicas y documentos institucionales, la investigación acredita que el mandato de supervisión humana impuesto por el Reglamento Europeo de Inteligencia Artificial resulta ineficaz en la práctica. La eficiencia algorítmica debilita la capacidad de decisión de los funcionarios, la independencia de los jueces y principios constitucionales clave. Por ello, se demuestra que es urgente crear una “reserva de humanidad” que permita devolver el análisis cualitativo y la ponderación empática a las decisiones públicas.

Palabras clave

Sesgo de automatización, inteligencia artificial, responsabilidad administrativa, justicia algorítmica, VioGén, RisCanvi

¹ Daniel Martínez Cristóbal. Universidad Rey Juan Carlos (URJC). Email: danielmcrystal@gmail.com
ORCID: <https://orcid.org/0000-0001-9754-5688>

Abstract

This article analyses, from an interdisciplinary socio-legal perspective, the impact of algorithmic risk assessment systems (VioGén and RisCanvi) on public administration and the judicial system in Spain. It examines how the use of these predictive tools has led professionals to disregard their own judgement due to “automation bias” and the pressure currently generated by defensive and precautionary dynamics within the civil service. By applying a qualitative approach that integrates legal sources, technical assessments, and institutional documents, the research establishes that the human oversight mandate imposed by the European Artificial Intelligence Act proves ineffective in practice. Algorithmic efficiency undermines the decision-making capacity of public officials, the independence of the judiciary, and key constitutional principles. Consequently, the study demonstrates the urgent need to establish a “reserve of humanity” to restore qualitative analysis and empathetic deliberation to public decision-making.

Key words

Automation bias, artificial intelligence, administrative procedure, algorithmic justice, VioGén, RisCanvi

Table of contents

1. Introducción	2237
2. El Estado algorítmico y la crisis de la discrecionalidad administrativa	2238
2.1. El sesgo de automatización (<i>automation bias</i>) en la función pública	2239
2.2. La ilusión del <i>Human in the Loop</i> (HITL) frente al Reglamento Europeo de IA	2239
3. La responsabilidad profesional y la presión de la administración.....	2240
3.1. El <i>street-level bureaucrat</i> frente a la dictadura del software	2240
3.2. El miedo a la responsabilidad y el auge de la Administración defensiva digital	2241
3.3. La reconfiguración de la deontología pública	2242
4. Estudio de VioGén y la gestión policial del riesgo	2242
4.1. Naturaleza de actuación, estructura y evolución normativa.....	2243
4.2. La valoración policial frente a la predicción estadística.....	2243
4.3. Impacto en la tutela de los derechos de las víctimas	2244
5. Estudio de RisCanvi y la determinación de actuación penitenciaria.....	2245
5.1. El algoritmo dentro de las Juntas de Tratamiento	2245
5.2. La independencia judicial frente a la autoridad algorítmica.....	2246
5.3. La colisión del modelo de reinserción social y la irrupción del derecho penal de autor.....	2247
6. Conclusiones	2248
Referencias	2249

1. Introducción

El proceso de transformación digital de las Administraciones públicas actuales ha pasado de la simple informatización y la gestión electrónica de expedientes hacia la incorporación de sistemas complejos de Inteligencia Artificial (IA) diseñados para ayudar o, directamente, para la toma automatizada de decisiones. Esta profunda evolución estructural y tecnológica ha dado lugar a lo que se conoce como el “Estado algorítmico”. En este nuevo escenario, la tecnología funciona como un elemento central que condiciona, organiza y, en múltiples supuestos prácticos, predetermina la resolución final que influye de manera directa en los derechos de los ciudadanos (Boix Palop 2020).

Esta investigación presenta el objetivo principal de cómo la forma de pensar de los trabajadores públicos y la propia organización administrativa responden ante esta delegación tecnológica. La investigación no se queda solo en un análisis teórico sobre cómo encaja la IA en la Constitución, sino que plantea un estudio de la autoridad, obediencia y responsabilidad que se generan cuando un funcionario debe contraponer su criterio profesional a una predicción matemática.

Para fundamentar sólidamente los argumentos jurídicos desarrollados a lo largo de este estudio y evitar caer en reflexiones puramente teóricas o especulativas que a menudo lastran los estudios sobre la aplicación legal de las nuevas tecnologías, esta investigación adopta una metodología sociojurídica de carácter interdisciplinar. Este enfoque se estructura en la triangulación del estudio de documentos y fuentes primarias generadas por las propias instituciones implicadas, la revisión de investigación académica especializada sobre el tema, y la integración de diferentes marcos teóricos. De esta forma, se unen los principios básicos del derecho constitucional y administrativo con los datos prácticos que aportan la sociología de las organizaciones y el estudio del comportamiento dentro de la administración pública.

El primer pilar de la base de este estudio procede del análisis de informes de auditoría y documentos de evaluación institucional recientes a partir de la revisión detallada de recientes informes de auditoría técnica de evaluación institucional. Estos textos permiten examinar con gran detalle el funcionamiento de las herramientas algorítmicas que son objeto de esta investigación.

Por lo que respecta al criticado sistema RisCanvi, que actualmente emplea el Departamento de Justicia de la Generalitat de Cataluña, se ha analizado en profundidad el contenido del *Informe Tiresias*, realizado en una auditoría independiente publicada en 2024, y que consta de un documento técnico sobre dicho algoritmo. De esta fuente original se han extraído métricas de rendimiento matemático y estadístico como las áreas bajo la curva AUC-ROC, la evidencia sobre los sesgos hacia determinados grupos de población y la constatación de las altas tasas de falsos negativos y falsos positivos.

De forma paralela, en el análisis del sistema estatal VioGén destaca el informe publicado por la organización *Eticas Research* en 2022, en el que se muestra de forma clara la presión administrativa sobre las dinámicas de asimetría, desigualdad y obediencia que surgen durante el contacto entre los agentes de policía, los formularios electrónicos cerrados para evaluar el riesgo (VPR y VPER) y las víctimas en sede policial.

La segunda línea metodológica que estructura este trabajo consiste en la revisión de la doctrina sobre la falta de justificación legal en las decisiones automatizadas adoptadas

por la Administración, así como de la incomprensible renuncia a la independencia por parte de los jueces de vigilancia penitenciaria frente a los cálculos algorítmicos, logrando analizar el impacto institucional y los daños que causa el mencionado sesgo de automatización.

Finalmente, para unificar la investigación de este estudio se ha apoyado en teorías sociológicas sobre el concepto de la “burocracia a nivel de calle”, combinándolo con otras teorías como la actual “burocracia defensiva” en el sector público. Este cruce de ideas permite ofrecer una visión completa y profunda sobre la manera en que la estructura administrativa del Estado puede comprometer las garantías del ciudadano frente a la automatización de las decisiones públicas en base a sistemas informáticos predictivos.

Gracias a este enfoque, es posible comprender cómo ciertas rutinas tecnológicas en la Administración pública pueden llegar a vulnerar derechos básicos en la práctica diaria, siempre bajo una apariencia de simple modernización informática y digitalización de los servicios del Estado dentro de esta distópica era digital.

2. El Estado algorítmico y la crisis de la discrecionalidad administrativa

La incorporación de herramientas algorítmicas predictivas en la Administración pública ha generado una fuerte crisis en la idea tradicional de la discrecionalidad administrativa, definida como el margen de libertad conferido al funcionario o autoridad para valorar las circunstancias particulares de un caso, buscando siempre el interés general y el principio de interdicción de la arbitrariedad de los poderes públicos, establecido en el artículo 9.3 de la Constitución Española (CE) (Bacigalupo 2023).

A pesar de ello, la búsqueda de lograr procesos más rápidos y de tratar todos los casos de manera uniforme ha fomentado el uso de sistemas que, materialmente, anulan la capacidad de valoración del empleado público (Ponce Solé 2024). La contradicción estructural del Estado algorítmico radica precisamente en su pretensión por neutralizar la vulnerabilidad, el error y la discrecionalidad humana, que termina consolidando una amenaza en forma de “arbitrariedad codificada”.

Estas herramientas de predicción actúan en la inmensa mayoría de las ocasiones como cajas negras (*black boxes*). Como están guiadas por modelos de aprendizaje automático (*machine learning*) o fórmulas matemáticas de alta complejidad, la relación causal y lógica entre los datos de entrada (*inputs*) y la recomendación de salida (*outputs*) resulta completamente incomprensible para la persona que tiene que firmar y validar la decisión final (Cerrillo i Martínez 2020).

Esta falta de transparencia choca directamente con el deber legal de justificar y motivar las resoluciones administrativas y judiciales como uno de los pilares del Estado de Derecho. Si un funcionario fundamenta una decisión basándose únicamente en el resultado de un algoritmo cuyo funcionamiento interno ignora o no puede comprender, la justificación jurídica se convierte en un argumento vacío, ya que la decisión se asume como correcta única y exclusivamente porque el sistema informático así lo ha calculado. Ante este escenario de indefensión de la ciudadanía emerge en la doctrina contemporánea la necesidad de reivindicar una “reserva de humanidad” en el diseño de los nuevos procedimientos administrativos (Ponce Solé 2019). Esta idea busca garantizar que aquellas decisiones públicas con un impacto grave en los derechos fundamentales

conserven, en todo momento, un espacio intocable para la reflexión moral, el análisis del contexto y la sensibilidad humana, elementos que ninguna máquina puede imitar.

2.1. *El sesgo de automatización (automation bias) en la función pública*

Para entender por qué los profesionales renuncian a su propio criterio frente a las máquinas, resulta imprescindible observar los datos que aportan la psicología cognitiva y la sociología de las organizaciones. El problema principal es el llamado “sesgo de automatización” (*automation bias*), mediante el cual las personas confían de manera incondicional e irreflexiva en las sugerencias, decisiones o predicciones generadas por sistemas automatizados. Al hacerlo, ignoran de forma sistemática cualquier evidencia contradictoria, incluso si esa información proviene de su propia experiencia profesional previa o de lo que observan en su entorno (Alon-Barkat y Busuioc 2023).

Los estudios científicos en materia de interacciones entre humanos y ordenadores (*Human-Computer Interaction*) advierten de forma unánime que este comportamiento no es fruto de la negligencia consciente o a errores de individuos aislados. En realidad, se trata de un mecanismo de respuesta neurológica frente al exceso de información y a la presión institucional contemporánea (Aguilar Ortiz 2026). En las Administraciones públicas donde hay que resolver un gran volumen de expedientes bajo estrictas métricas de rendimiento y productividad temporal, el sistema algorítmico se percibe inmediatamente como una vía rápida y segura, un verdadero atajo cognitivo, para resolver la constante incertidumbre que caracteriza la toma de decisiones.

Tanto en el sector público como en el privado, este fenómeno se traduce en que los trabajadores tienden a aceptar lo que recomienda el algoritmo entre el 70% y el 90% de las ocasiones, incluso cuando el sistema incurre en errores evidentes o proporciona directrices absurdas (Vorlux AI 2024). El sesgo de automatización empeora porque la sociedad tiende a creer que los programas basados en números y modelos estadísticos tienden a la objetividad matemática e infalibilidad.

Esta subordinación a la tecnología se ha documentado en múltiples facetas en el mundo actual, llegando a acuñarse el término de “muertes por GPS” (Bridle 2020) para describir aquellos supuestos extremos en los que individuos perfectamente racionales han seguido a rajatabla instrucciones de navegación erróneas, ignorando las señales visuales del mundo físico, hasta precipitarse a un acantilado o adentrarse en un desierto. Cuando esto ocurre en el sector público, los programas que debían ser solo sistemas de asistencia o apoyo se convierten en los verdaderos responsables de las decisiones, anulando así todos los mecanismos de protección que el ordenamiento jurídico preveía para tutelar al ciudadano frente al poder del Estado.

2.2. *La ilusión del Human in the Loop (HITL) frente al Reglamento Europeo de IA*

Consciente de los inmensos riesgos que entraña la delegación de decisiones críticas en modelos algorítmicos, el legislador europeo ha pretendido articular una serie de mecanismos para prevenir o minimizar los riesgos para los derechos fundamentales, exigiendo que la persona física encargada de la supervisión sea plenamente consciente del riesgo del ya mencionado “sesgo de automatización”.

El Reglamento Europeo de Inteligencia Artificial (*AI Act*, Reglamento UE 2024/1689) establece explícitamente en su artículo 14.4.b la obligación de que todo sistema de IA considerado como de alto riesgo sea diseñado y desarrollado para permitir una supervisión humana efectiva durante todo el período que esté en uso y decida, si las circunstancias lo exigen, apartarse del dictamen de la máquina (Unión Europea 2024). Este sistema, conocido como *Human in the Loop* (HITL), tiene como propósito exigir que los encargados de vigilar la máquina entiendan sus límites y sepan que existe el riesgo del sesgo de automatización.

No obstante, el sistema está configurado como un sofisticado cortafuegos de responsabilidad legal, denominado *rubber-stamping* (Alon-Barkat y Busuioc 2023). La teoría idealizada del HITL que subyace en el Reglamento presupone la existencia de un trabajador que dispone del tiempo necesario para revisar cada expediente y posee la competencia técnica e informática para interpretar complejos modelos de probabilidad y que posee el respaldo para atreverse a contradecir a un sistema informático que ha costado millones de euros (Revista de Derecho PUCP 2024).

La situación real de las Administraciones españolas, marcada por la escasez de recursos materiales y humanos, dista enormemente de este modelo. La revisión rápida se evidencia cuando se observan, de forma reiterada, validaciones sistemáticas de informes sin ningún tipo de comentario; no se analizan los expedientes a fondo y el trabajador va perdiendo poco a poco su capacidad de pensar de forma crítica, entrando en una progresiva obsolescencia cognitiva, al delegar su pericia analítica de forma ininterrumpida en la inteligencia artificial (Aguilar Ortiz 2026).

Si el marco institucional y el diseño organizativo de la función pública no fomentan y protegen la discrepancia fundamentada frente al algoritmo el modelo HITL termina actuando como una ficción legitimadora. La intervención humana reviste de validez jurídica el dictamen técnico, confiriéndole una plena presunción de acierto legal. Al mismo tiempo, el sistema transfiere la totalidad de la carga de la responsabilidad legal, penal y disciplinaria a un trabajador víctima del diseño sistémico y del sesgo de automatización, ha sido relegado a la condición de mero validador pasivo.

3. La responsabilidad profesional y la presión de la administración

La conexión que se produce entre la estructura institucional jerarquizada y el uso de tecnologías de predicción demuestra que los problemas del Estado algorítmico no son de naturaleza puramente técnica ni se resuelven exclusivamente con un mejor código fuente. En realidad, nos encontramos ante un fenómeno organizativo y sociopolítico que transforma las relaciones de poder en la llamada “burocracia a nivel de calle” (*street-level bureaucracy*), donde es necesario entender las causas subyacentes que inducen al empleado público a ceder voluntariamente su poder de decisión en favor del sistema algorítmico.

3.1. El street-level bureaucrat frente a la dictadura del software

La teoría de la “burocracia a nivel de calle” (Lipsky 1980/2010) defiende que la verdadera aplicación de las políticas públicas depende de los trabajadores de la Administración que están en primera línea (policías, asistentes sociales, educadores, personal sanitario o

funcionarios de prisiones) y que interactúan diariamente de forma directa con los ciudadanos.

Estos profesionales siempre han trabajado en situaciones caracterizadas por la ambigüedad normativa y la escasez crónica de recursos (Lipsky 1980/2010). Para sobrevivir a la enorme presión de sus obligaciones, han desarrollado rutinas de trabajo y han utilizado su libertad de decisión para adaptar o flexibilizar las directrices políticas a la compleja realidad específica de las situaciones individuales que se presentaban ante ellos.

Con la llegada de los programas informáticos predictivos, esa libertad de decisión que antes residía en el contacto humano, empático y directo de los empleados públicos de base se traslada hacia los programadores informáticos, los expertos en datos y los directivos que diseñan los sistemas de software y fijan los niveles de riesgo (*system-level bureaucracy*).

El sistema algorítmico opera como un supervisor que estandariza la interacción, haciendo que el trato con el ciudadano sea igual para todos, lo que obliga a encajar todas las situaciones humanas dentro de los estrechos márgenes de un formulario digital compuesto por casillas de verificación. Así, el trabajador público queda relegado a la condición de recolector de datos (*data gatherer*), cuya función principal es introducir adecuadamente en el sistema los *inputs* necesarios para que la IA realice la función de procesamiento cognitivo y toma de decisión, lo que le hace perder la oportunidad de comprender y empatizar con la persona que tiene delante (Alon-Barkat y Busuioc 2023).

3.2. El miedo a la responsabilidad y el auge de la Administración defensiva digital

La actuación de los empleados públicos está condicionada por el rechazo al riesgo, el miedo a equivocarse y la necesidad de proteger su carrera profesional. El uso generalizado de algoritmos ha propiciado la aparición de lo que se denomina “burocracia defensiva” o “Administración defensiva digital” (Boto 2023, Ponce Solé 2024). Este concepto explica cómo los empleados evitan tomar decisiones propias o novedosas por miedo a enfrentarse a castigos disciplinarios, económicos o penales.

El desequilibrio al asumir los riesgos es la causa principal de esta actitud conservadora, por la que, psicológicamente, cuando el trabajador se encuentra los riesgos frente a su pantalla, la conclusión es que resulta mucho más favorable obedecer siempre las instrucciones del sistema algorítmico. Si un trabajador público, basándose en su experiencia y en el análisis personal del caso concreto, decide contradecir la propuesta de la inteligencia artificial, asume en solitario la obligación de la carga probatoria y argumentativa. Para ello, deberá redactar informes detallados que justifiquen su decisión ante sus superiores.

El problema es mayor si, tras aplicar su propio criterio profesional ponderado, empático y reflexivo, se produce un fallo inevitable (por ejemplo, si un recluso al que se le concedió el permiso termina delinquiendo, o si un maltratador al que se le bajó el nivel de alerta termina atacando de nuevo a la víctima). En esos supuestos, toda la responsabilidad disciplinaria, legal, mediática y social recaerá directa y exclusivamente sobre él por haber ignorado un cálculo estadístico oficial. Tanto la sociedad como la Administración le

acusarán de negligencia profesional por haber ignorado de manera injustificada las advertencias estadísticas que proporcionó el sistema informático.

Por el contrario, la situación opuesta favorece la falta de responsabilidad, ya que si el trabajador obedece al programa informático y la máquina se equivoca provocando daños con consecuencias fatales, el error ya no se considera un fallo humano punible, sino que se asume como un problema del protocolo o un margen de error estadístico propio de la tecnología.

Ante esta situación, disentir y contradecir al algoritmo supone un riesgo enorme para el futuro laboral del profesional (Ponce Solé 2024). El tradicional “miedo a la firma” de la Administración clásica pasa intacto al entorno digital. Esto provoca que el empleado público prefiera protegerse detrás de la presunta fiabilidad del sistema informático, en lugar de ejercer los principios que le atribuye la Constitución (Bueno 2024).

3.3. La reconfiguración de la deontología pública

A medio y largo plazo, esta situación institucional provoca un cambio importante en los valores éticos que deben guiar el servicio público. El artículo 103 de la Constitución Española establece que la Administración Pública sirve con objetividad los intereses generales y actúa de acuerdo con los principios de eficacia, jerarquía, descentralización, desconcentración y coordinación, con sometimiento pleno a la ley y al derecho. Con ello, se exige a los trabajadores públicos un compromiso moral y ético para conseguir un trato justo y buscar la verdadera equidad a la hora de resolver cada caso concreto.

Con la instauración silenciosa de la gobernanza algorítmica, la gestión administrativa ha apartado ese compromiso ético tradicional y ha puesto en su lugar la simple obediencia a las reglas y a los sistemas informáticos (San Martín Segura 2023). De este modo, la profesionalidad de un empleado público ya no se evalúa por su inteligencia para analizar y solucionar problemas en base a un profundo criterio jurídico.

En su lugar, el desempeño se mide por la rapidez al procesar datos y por cumplir los protocolos sin causar retrasos. Esta nueva forma de trabajar genera una rutina de trabajo sin espíritu crítico y destruye el trato directo con el ciudadano. Como resultado, este deja de ser visto como una persona con derechos propios para convertirse, a los ojos de la institución, para convertirse en un conjunto de variables numéricas (*inputs*) que debe introducir en el ordenador para que el programa genere el correspondiente acto administrativo (*output*).

4. Estudio de VioGén y la gestión policial del riesgo

Para ilustrar estas complejas dinámicas sociojurídicas es necesario analizar el Sistema de Seguimiento Integral en los casos de Violencia de Género (VioGén), dependiente de la Secretaría de Estado de Seguridad del Ministerio del Interior. VioGén es uno de los sistemas de prevención policial y evaluación de riesgos delictivos más grandes de Europa. Este programa combina normas legales y fórmulas estadísticas para predecir si un agresor volverá a actuar y, en función de ello, asignar protección a las mujeres afectadas (López-Ossorio *et al.* 2019, San Martín Segura 2023).

4.1. Naturaleza de actuación, estructura y evolución normativa

VioGén es un sistema predictivo que funciona aplicando métodos matemáticos que calculan la amenaza basándose en un amplio historial de datos estadísticos de criminalidad. Su funcionamiento se articula a través de la recogida de información mediante dos formularios por parte de los policías encargados del caso. El primero es la Valoración Policial del Riesgo (VPR), que se completa en el mismo momento en que la mujer presenta la denuncia. Posteriormente, se completará la Valoración Policial de la Evolución del Riesgo (VPER) para revisar y poner al día el nivel de alerta durante toda la fase de seguimiento y vigilancia protectora (López-Ossorio *et al.* 2019, Ministerio del Interior 2025).

La versión que se utiliza hoy en día (VPR5) cruza y analiza alrededor de 43 factores de riesgo, entre los que se incluyen el historial delictivo del agresor, la situación laboral de la víctima, circunstancias en las que ocurrió el maltrato, episodios previos de violencia, celotipia, consumo de tóxicos, etc. Una vez procesada toda esta información, el programa informático estima la probabilidad estadística de que se produzca un nuevo episodio de violencia (reincidencia o revictimización), clasificada en cinco niveles (“No apreciado”, “Bajo”, “Medio”, “Alto” y “Extremo”).

Cada uno de estos niveles lleva aparejado la adopción de una serie de medidas de protección policial concretas, como visitas periódicas o asignación de patrullas de vigilancia, tal y como establece la Instrucción 1/2025 de la Secretaría de Estado de Seguridad (Academia Geopol 2025). Por este motivo, el resultado matemático que muestra VioGén determina y condiciona de forma automática la actuación de las Fuerzas y Cuerpos de Seguridad del Estado frente al agresor y la víctima.

4.2. La valoración policial frente a la predicción estadística

Las instituciones defienden que el sistema VioGén es plenamente garantista porque cuenta con una supervisión humana real (*Human in the loop*) debido a la discrecionalidad policial para modificar o alterar el nivel de amenaza que calcula el algoritmo (San Martín Segura 2023), al igual que ocurre con protocolos autonómicos paralelos como el QPVR en Cataluña o el EPV-R en el País Vasco.

Sin embargo, esta libertad de modificación del criterio algorítmico (*override*) está limitada en la práctica, ya que el agente policial solo está autorizado para elevar el nivel de peligro, y en ningún caso a la baja (Éticas Research and Consulting 2022). Esta opción se pensó para que el policía pudiera aumentar la protección si, en base a su intuición o a su experiencia criminológica, observase señales de peligro que el cuestionario cerrado del algoritmo no ha logrado registrar.

A pesar de ello, los estudios demuestran que esta discrecionalidad para corregir a la máquina es irrelevante y marginal. Las auditorías externas e independientes realizadas sobre el sistema y los informes en las comisarías señalan que las fuerzas policiales apenas han modificado alrededor del 5% de las valoraciones de riesgo generadas por la inteligencia artificial de VioGén (Éticas Research and Consulting 2022).

Esta sumisión a la máquina se explica, en gran parte, por la sobrecarga de trabajo que padecen las unidades especializadas. Además, el propio diseño del programa exige a los policías redactar informes adicionales para justificar por qué no están de acuerdo con el

algoritmo (Éticas Research and Consulting 2022). Ante esta presión en el trabajo diario, el agente de policía termina aceptando el nivel de riesgo evaluado por el sistema de forma automática, cediendo su criterio profesional y su empatía protectora a la estandarización predictiva.

4.3. Impacto en la tutela de los derechos de las víctimas

Delegar el cálculo del riesgo en manos de un sistema tan sensible como VioGén conlleva consecuencias directas y en ocasiones irreversibles en la tutela efectiva de los derechos fundamentales de las víctimas de violencia de género. La confianza en el cálculo estadístico ha destapado vulnerabilidades críticas directamente relacionadas con el diseño del contacto y trato con la víctima en un entorno de alto estrés.

El protocolo exige que los formularios VPR y VPER sean cumplimentados mediante respuestas cerradas a la persona afectada obtenidas en el momento de la interposición de la primera denuncia en la comisaría, una situación límite en la que casi todas se encuentran en estado de shock emocional o sufren un fuerte trauma. Las diferentes auditorías independientes comprueban que más del 80% de las víctimas entrevistadas se hallan en un estado de *shock* emocional, disociación o terror absoluto por el que tienen muchas dificultades para entender el lenguaje y el significado exacto de las preguntas planteadas por el formulario automatizado de VioGén, lo que provoca que las respuestas recogidas no sean del todo precisas (Éticas Research and Consulting 2022). Estas dificultades aumentan de manera exponencial en aquellos casos donde inciden otros factores, como barreras idiomáticas, contextos culturales distintos o bajos niveles formativos.

Al introducir datos poco exactos o incompletos, el algoritmo no consigue detectar factores humanos y sociales que son claves para saber si existe realmente un peligro inminente. Esto genera la proliferación de “falsos negativos” algorítmicos, casos donde el programa establece que no hay riesgo cuando, en realidad, la amenaza es muy alta. Al igual que se ha documentado en otros países que usan sistemas parecidos, como el protocolo DASH de la policía en el Reino Unido (diversas investigaciones concluyeron que hasta el 67% de los casos que acabaron en muerte violenta habían sido previamente clasificados por el sistema como casos de riesgo bajo o sin peligro), VioGén adolece de puntos ciegos estructurales inherentes a su modelo matemático predictivo (Éticas Research and Consulting 2022).

En España, el programa VioGén archiva de forma automática aquellos expedientes que pasan cierto tiempo sin registrar nuevas incidencias ni denuncias, pasándolos a la lista de “casos inactivos”. Tras el repunte de asesinatos por violencia de género en los últimos años, los informes han demostrado que una parte importante de los crímenes se cometió contra mujeres cuyos casos el algoritmo había clasificado como inactivos o sin riesgo (Éticas Research and Consulting 2022).

En lugar de aprovechar estos dolorosos fracasos algorítmicos para poner en duda la metodología y los principios originarios del sistema, la solución del Ministerio ha consistido en agregar parches tecnológicos, como ampliar los plazos antes de archivar un caso dentro de las bases de datos de VioGén e incrementar la dotación de recursos telemáticos pasivos como la compra de más pulseras telemáticas. En cambio, se ha

dejado de lado la necesidad de potenciar radicalmente la formación específica de los agentes y de devolverles el protagonismo a la hora de reconocer la amenaza y el riesgo.

5. Estudio de RisCanvi y la determinación de actuación penitenciaria

El programa RisCanvi, creado y utilizado por el Departamento de Justicia de la Generalitat de Cataluña desde el año 2009, ilustra la regulación del artículo 25.2 de la Constitución Española, por el que las penas y medidas de seguridad en los ámbitos administrativo-penitenciario y judicial deben estar orientadas a la reeducación y reinserción social (Andrés-Pueyo 2025, Jiménez Arandía 2026).

RisCanvi es un protocolo tecnológico de evaluación del riesgo diseñado para predecir y evaluar la peligrosidad futura de los reclusos y, consecuentemente, orientar las posteriores decisiones sobre la evolución de los grados penitenciarios, la concesión de permisos de salida ordinarios y la aprobación de libertades condicionales (Andrés-Pueyo 2025, Jiménez Arandía 2026).

5.1. El algoritmo dentro de las Juntas de Tratamiento

RisCanvi estructura su predicción para evaluar el riesgo a través de un programa informático integrado en el sistema penitenciario en Cataluña, en el que el algoritmo procesa decenas de factores de riesgo individuales e historiales delictivos, entre fijos (como los antecedentes o la edad al cometer el primer delito) y otros cambiantes (como la conducta diaria en la cárcel). Estas valoraciones están orientadas a calcular la probabilidad de que ocurran cinco tipos específicos de violencia o incidencias en el futuro del interno, como es el riesgo de violencia autodirigida (suicidio o autolesiones), violencia interna (agresiones a funcionarios u otros reclusos), quebrantamiento de condena (fugas), reincidencia general (comisión de cualquier nuevo delito) y reincidencia violenta (comisión de delitos violentos o contra las personas) (Andrés-Pueyo 2025).

Los equipos multidisciplinares de trabajadores de las prisiones, que componen las Juntas de Tratamiento, se basan frecuentemente en estas valoraciones algorítmicas para apoyar sus propuestas a la hora de otorgar una flexibilización penitenciaria con relación a permisos, conceder el tercer grado o pedir la libertad condicional.

A pesar de este uso generalizado, una auditoría realizada por la empresa Dribia Data Research sobre el algoritmo (el *Informe Tiresias*, elaborado y publicado en enero de 2024) ha revelado fallos importantes sobre el funcionamiento del sistema. El software predictivo adolece de deficiencias estructurales, metodológicas, criminológicas y operativas que inciden en las tasas de precisión real del sistema (Dribia Data Research 2024).

Desde el punto de vista informático, el sistema actual trabaja con un código anticuado de cuya evaluación dependen los años de libertad de miles de personas, se sustenta en un código anticuado y frágil (integrado precariamente en el lenguaje propio de consultas de base de datos psql). Además, no cuenta con manuales actualizados ni un control de versiones, y es necesario la contratación de informáticos externos a la Administración para poder realizar la supervisión (Dribia Data Research 2024).

Al analizar su diseño, los auditores descubrieron que el programa fue calibrado prioritariamente por la Administración para mantener unas tasas bajas de Falsos Positivos con el objetivo de no catalogar erróneamente de peligrosos a internos que no lo son (FPR del 1% al 31%,). Por ello, el algoritmo equilibra los números cometiendo el error contrario e incurre en una infraestimación generalizada del riesgo real. Esto significa que la configuración produce una cantidad enorme de Falsos Negativos, al predecir que un interno no supone riesgo cuando en realidad sí acaba delinquiendo o agrediendo, con fallos que van del 40% al 100%, dependiendo de la medida que se use (Dribia Data Research 2024).

El ejemplo más claro de este fallo se encontró en la fase de cribado (*screening*) para predecir el riesgo de quebrantamiento de condena. En este caso, las cifras reales de la auditoría mostraron que el programa tuvo una Tasa de Verdaderos Positivos (TPR) del 0 % y una inasumible Tasa de Falsos Negativos del 100 % (Dribia Data Research 2024), cifras que apenas disminuyen en la evaluación completa (44 % y 56 %, respectivamente). Esta ineficacia para advertir el riesgo real se reproduce de forma sistémica en categorías de extrema gravedad, donde en la predicción de reincidencia violenta el sistema pasa por alto al 84,25 % de los reclusos de riesgo en fase de cribado y al 72,29 % en su fase completa. Patrones similares se observan en la violencia autoinfligida (88,87 % de Falsos Negativos iniciales) y en la violencia en el centro de internamiento (71,71 %). Paradójicamente, en la predicción de reincidencia general, el paso del cribado a la escala completa desploma el acierto predictivo del 60,03 % al 18,72 %, reafirmando una inseguridad jurídica en la valoración penitenciaria.

Estos datos demuestran que la evaluación de amenaza que determina el sistema adolece de graves fallos a la hora de predecir determinados contextos de riesgo penitenciario (Jiménez Arandía 2024). A pesar de la evidencia de estos errores, los resultados siguen utilizándose como la guía principal para la toma de decisión con relación a la concesión de beneficios y mejoras en la condena de los presos en Cataluña.

5.2. La independencia judicial frente a la autoridad algorítmica

El peso de RisCanvi trasciende las barreras de la prisión y se introduce en el ámbito judicial de la mano de los Jueces de Vigilancia Penitenciaria, que tienen el deber de proteger los derechos fundamentales de los internos frente a las decisiones administrativas-penitenciarias, para evaluar los informes manifestados por las Juntas de Tratamiento sobre las peticiones de libertad o permisos, en los que la recomendación de riesgo valorada por el algoritmo figura de manera destacada (Jiménez Arandía 2026).

Esto crea el dilema entre la teórica independencia judicial y la aceptación de la nueva “autoridad sagrada de la técnica” donde se supone que la máquina informática siempre tiene la razón (Gómez Colomer 2023). Aunque el juez ostenta la potestad jurisdiccional para apartarse del informe penitenciario y desoír sus recomendaciones, atreverse a contradecir a un sistema de riesgo automatizado que ha clasificado a un individuo con un riesgo “Alto” requiere una motivación y un alto nivel de certidumbre subjetiva además de una enorme presión por si algo sale mal (Alemán Aróstegui 2023).

El miedo a las consecuencias de equivocarse al firmar la concesión de un beneficio penitenciario en contra del criterio del algoritmo (si el recluso reincidiera cometiendo un

nuevo delito violento durante el disfrute de dicho permiso) empuja al juez a ser prudente y a darle siempre la razón al cálculo de la máquina (Gómez Colomer 2023).

Resulta casi imposible entender las estadísticas y factores que pondera el programa debido a la opacidad técnica y al secreto de sus variables de código, lo que impide que la autoridad judicial pueda llevar a cabo un control de forma exhaustiva (Alemán Aróstegui 2023). Como consecuencia de esto, el juez asume una actitud de sumisión y deferencia casi absoluta hacia la ponderación del algoritmo, externalizando su responsabilidad en la inteligencia artificial. De esta manera, la vital función de los Juzgados de Vigilancia Penitenciaria se desvirtúa, y el análisis individual de cada caso se transforma en una simple validación, dando por hecho que la estadística del sistema nunca falla (Alemán Aróstegui 2023).

5.3. La colisión del modelo de reinserción social y la irrupción del derecho penal de autor

Esta obediencia ciega al algoritmo choca de frente con el artículo 25.2 de la Constitución Española, que establece como fin primordial de las penas privativas de libertad la reeducación y la reinserción social efectiva del delincuente (Solar Calvo y Lacal Cuenca 2021). Para que esta meta constitucional se cumpla, cualquier mejora en la condena debe basarse en valorar la evolución real del interno, evaluar el esfuerzo real de la persona, analizar los cambios individuales a través del arrepentimiento, y el desarrollo conductual diario en prisión.

En contraste, el sistema RisCanvi fundamenta su evaluación en una concepción estrictamente estática, determinista y retrospectiva del individuo. Asimismo, la auditoría del *Informe Tiresias* ha constatado la existencia de sesgos algorítmicos discriminatorios preocupantes que atentan contra la igualdad (Dribia Data Research 2024). El algoritmo presenta un margen de error significativamente superior y produce una mayor proporción de resultados adversos injustificados sobre colectivos vulnerables o minoritarios. Esta disfunción afecta de manera desproporcionada a la población reclusa extranjera, a las mujeres internas (especialmente en la evaluación errónea del riesgo vital de suicidio o violencia autodirigida, donde son castigadas por falsos negativos que les niegan la ayuda clínica necesaria) y los internos menores de 30 años, reproduciendo de este modo los sesgos estructurales presentes en los propios datos históricos con los que fue entrenada el programa (Jiménez Arandía 2026).

Al fundamentar la concesión o denegación de los beneficios penitenciarios y la libertad en correlaciones estadísticas grupales predeterminadas (donde el interno no es evaluado por su evolución individual ni por lo que ha hecho en prisión, sino por la conducta histórica de su perfil demográfico), el sistema penitenciario y judicial español juzga a la persona por lo que “es” y no por lo que “ha hecho”, asimilando postulados propios del derecho penal de autor, legitimándolos bajo una apariencia de objetividad cibernética e infalibilidad técnica (Alemán Aróstegui 2023).

Bajo la tiranía del Estado algorítmico, el interno deja de ser considerado por las instituciones como un sujeto de derecho inmerso en un proceso de reinserción reduciéndolo a un mero individuo de variables inmutables de riesgo. En consecuencia, la expectativa a la progresión de grado penitenciario queda bloqueada, no por su conducta en la prisión, sino por el infortunio de pertenecer a un perfil demográfico que

es penalizado automáticamente por la automatización del código (Jiménez Arandía 2026).

6. Conclusiones

La progresiva integración de la inteligencia artificial predictiva en la gestión pública de los riesgos sociales, criminológicos y delictivos en el ordenamiento jurídico español ofrece una evidencia clara. La búsqueda de la eficiencia tecnológica, la aparente neutralidad de los algoritmos y la velocidad procesal y la rapidez que promete la instauración irreversible del Estado algorítmico, conllevan un importante coste social y jurídico. Este fenómeno genera una opacidad estructural que desplaza paulatinamente el juicio ético y la indispensable ponderación humana frente a la vulnerabilidad ciudadana.

El análisis de la implementación de sistemas como VioGén y RisCanvi desmonta la idea que concibe a los algoritmos como herramientas neutrales y objetivas. El estudio demuestra que el diseño del código, los sesgos ocultos, los criterios económicos y los márgenes de error estadístico no actúan como un simple apoyo técnico al servicio del bienestar común. Al contrario, la IA funciona hoy como una estructura que genera asimetrías de poder, capaz de alterar en la práctica los principios éticos y la capacidad de decisión de los empleados públicos, eludiendo los mecanismos de control habituales.

Como consecuencia directa de esta opacidad técnica, el mecanismo de supervisión humana (*Human in the Loop*), diseñado para proteger los derechos previstos por normativas como el Reglamento Europeo de Inteligencia Artificial (*AI Act*), corre el riesgo de resultar ineficaz. En la práctica policial, administrativa y judicial, este control puede quedar reducido a un trámite formal, orientado a legitimar el sistema y a salvaguardar la apariencia de legalidad. Bajo la falsa seguridad de una supervisión humana garantista del proceso, se consolida una delegación casi total de las competencias coercitivas del Estado en modelos matemáticos opacos. Dicha delegación contraviene la prohibición de la arbitrariedad de los poderes públicos y el deber ineludible de motivación de los actos administrativos, toda vez que los fundamentos probabilísticos de la decisión resultan incomprensibles tanto para el funcionario encargado de su convalidación como para el juez que asume el dictamen y para el ciudadano que padece sus efectos limitativos de derechos.

En este contexto institucional, el sesgo de automatización no puede seguir concibiéndose como un simple error cognitivo ocasional o una vulnerabilidad psicológica del trabajador público. Por el contrario, es el resultado directo y previsible de un diseño organizativo problemático. Este entorno, condicionado por la “Administración defensiva digital”, somete al empleado público a una fuerte presión, obligándolo a tomar decisiones bajo una tensión constante.

Dentro de este entorno ejercer el deber ético de cuestionar la recomendación del algoritmo requiere que el trabajador público asuma un riesgo profesional que resulta casi inasumible para la mayoría de los trabajadores, ya que se enfrenta a la sobrecarga diaria de tareas y al temor a sufrir sanciones disciplinarias, exposición mediática o consecuencias penales que afecten gravemente a su carrera, en caso de apartarse del dictamen automatizado si este resultara, a posteriori, ser erróneo.

Con relación al sistema VioGén, la subordinación del agente policial al criterio del sistema provoca una estandarización inflexible de las complejas situaciones que caracterizan la violencia de género. La rígida estructura algorítmica propicia la aparición de umbrales de falsos negativos, dejando en situación de desprotección a víctimas cuyo nivel de riesgo real es invisibilizado por el sistema. Estos errores se materializan cuando el profesional, presionado por el volumen de trabajo y limitado por la rigidez de los protocolos, no percibe a tiempo el peligro o no se atreve a contradecir la estadística del ordenador, claudicando ante la presunta infalibilidad estadística.

Una conclusión similar se extrae de la aplicación del sistema RisCanvi en Cataluña. La autoridad concedida en el día a día a la predicción del ordenador anula la capacidad de razonamiento y la independencia que la ley exige al Juez de Vigilancia Penitenciaria. La denegación sistemática de beneficios penitenciarios fundamentada en dictámenes algorítmicos sesgados vulnera el principio de reinserción y el tratamiento individualizado de la condena, establecido en el artículo 25.2 de la Constitución Española.

De esta forma, la justicia predictiva sustituye la valoración del comportamiento y el progreso individual del recluso por un modelo estadístico rígido que ha demostrado tener sesgos importantes, penalizando de forma discriminatoria a grupos vulnerables, como la población extranjera, las minorías o los internos jóvenes, quienes carecen de medios para defenderse frente a una predicción basada en datos anteriores de otras personas.

Ello supone una regresión hacia los postulados del derecho penal de autor, por el cual el individuo deja de ser juzgado por sus actos para ser condenado por su pertenencia a un perfil estadístico. Para proteger a los ciudadanos frente al uso inadecuado de la inteligencia artificial, resulta urgente establecer una reserva de humanidad dentro de la Administración Pública y en los procedimientos administrativos y judiciales, garantizando los derechos fundamentales frente a la automatización del sistema.

Es imprescindible organizar una protección de la Administración para preservar la capacidad y libertad de decisión del empleado público sobre la discrepancia fundamentada frente al dictamen del algoritmo, de manera que se asegure el tiempo necesario para analizar cada caso. Además es imperativo recibir la formación necesaria para comprender el funcionamiento interno de estos sistemas, resultando plenamente transparentes y comprensibles para la sociedad.

Todo este conjunto de medidas legales y técnicas es fundamental para proteger el espacio destinado a la empatía, a la duda razonable y al juicio ético como única vía para que el Estado supere la tentación de buscar una eficiencia basada en la obediencia ciega a las decisiones automatizadas.

Referencias

- Academia Geopol, 2025. *Instrucción 1/2025 de la Secretaría de Estado de Seguridad por la que se establece un nuevo protocolo para la valoración policial del nivel de riesgo de violencia de género (VioGén-2)*.
- Aguilar Ortiz, M., 2026. *Human in the Loop y aprendizaje continuo: la competencia crítica para supervisar IA* [en línea]. OpenWebinars. Disponible en:

<https://openwebinars.net/blog/human-in-the-loop-y-aprendizaje-continuo-la-competencia-critica-para-supervisar-ia/>

- Alemán Aróstegui, L., 2023. El uso de RISCANVI en la toma de decisiones penitenciarias. *Estudios Penales y Criminológicos* [en línea], 44(Ext.), 1-43. Disponible en: <https://doi.org/10.15304/epc.44.8884>
- Alon-Barkat, S., y Busuioc, M., 2023. Human–AI interactions in public sector decision making: “automation bias” and “selective adherence” to algorithmic advice. *Journal of Public Administration Research and Theory* [en línea], 33(1), 153-169. Disponible en: <https://doi.org/10.1093/jopart/muac007>
- Andrés-Pueyo, A., 2025. El RisCanvi: entre la predicción algorítmica y la decisión profesional para la prevención de la reincidencia penitenciaria. *ReCRIM: Revista de l'Institut Universitari d'Investigació en Criminologia i Ciències Penals de la UV* [en línea], número extra 25, 21-35. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=10504736>
- Azuaje Pirela, M. C., et al., 2025. Intervención, supervisión y control humano en las decisiones automatizadas. *Derecho PUCP* [en línea], (95), 189–226. Disponible en: <https://doi.org/10.18800/derechopucp.202502.006>
- Bacigalupo, M., 2023. La vinculación de la administración pública a la ley y al derecho. En: F. Velasco y M. Darnaculleta, eds., *Manual de derecho administrativo*. Madrid: Marcial Pons.
- Boix Palop, A., 2020. 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* [en línea], 1, 223-269. Disponible en: https://doi.org/10.37417/RPD/vol_1_2020_33
- Boto Álvarez, A., 2023. Nuevas barreras burocráticas: la Administración defensiva digital. *Documentación Administrativa*, (10), 40–58. Disponible en: <https://doi.org/10.24965/da.11200>
- Bridle, J., 2020. *La nueva edad oscura: La tecnología y el fin del futuro*. Madrid: Debate.
- Bueno Armijo, A., 2024. *La burocracia defensiva* [en línea]. Madrid: Marcial Pons. Disponible en: <https://marcialpons-openaccess.es/index.php/juridicas/catalog/download/burocracia-defensiva/44/912?inline=1>
- Cerrillo i Martínez, A., 2020. ¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos? *European Review of Digital Administration & Law* [en línea], 1(1-2), 17-XX. Disponible en: <https://www.erdalreview.eu/free-download/97888255389603.pdf>
- Dribia Data Research, 2024. *Informe Tiresias: Auditoria de l'algorisme RisCanvi (Versió 9 de gener de 2024)* [en línea]. Generalitat de Catalunya, Departament de Justícia. Disponible en: <https://repositori.justicia.gencat.cat/bitstream/handle/20.500.14226/1321/auditoria-algorisme-riscanvi-informe-final.pdf>

- Éticas Research and Consulting, 2022. *Auditoría de sesgos e incertidumbres algorítmicas en torno a las violencias de género. El caso del sistema VioGén.*
- Gómez Colomer, J. L., 2023. *El juez robot: la independencia judicial en peligro.* Valencia: Tirant lo Blanch.
- González Jaspe, J., 2024. EU AI Act artículo 14: Supervisión humana — Qué debe implementar tu empresa antes de agosto 2026. *Vorlux AI* [en línea], 25 de febrero. Disponible en: <https://vorluxai.com/es/blog/ley-ia-europea-supervision-humana-articulo-14/>
- Jiménez Arandía, P., 2024. RisCanvi: el misterioso algoritmo que evalúa a los presos en Cataluña de forma “azarosa”. *El País* [en línea], 3 de julio. Disponible en: <https://elpais.com/expres/2024-07-03/riscanvi-el-misterioso-algoritmo-que-evalua-a-los-presos-en-cataluna-de-forma-azarosa.html>
- Jiménez Arandía, P., 2026. Un algoritmo dificulta la reinserción de los presos vulnerables en Catalunya. *Público* [en línea], 6 de marzo. Disponible en: <https://www.publico.es/politica/algoritmo-dificulta-reinsercion-presos-vulnerables-catalunya.html>
- Lipsky, M., 2010. *Street-level bureaucracy: Dilemmas of the individual in public services.* Nueva York: Russell Sage Foundation. (Obra original publicada en 1980).
- López-Ossorio, J. J., Loinaz, I., y González-Álvarez, J. L., 2019. Protocolo para la valoración policial del riesgo de violencia de género (VPR4.0): revisión de su funcionamiento. *Revista Española de Medicina Legal* [en línea], 45(2), 52-58. Disponible en: <https://doi.org/10.1016/j.reml.2019.01.002>
- Ministerio del Interior (Secretaría de Estado de Seguridad), 2025. *La valoración policial del riesgo de violencia contra la mujer pareja en España (Sistema VioGén).*
- Ponce Solé, J., 2019. Inteligencia artificial, derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico. *Revista General de Derecho Administrativo*, 50.
- Ponce Solé, J., 2024. Inteligencia Artificial, decisiones administrativas discrecionales totalmente automatizadas y alcance del control judicial: ¿indiferencia, insuficiencia o deferencia? *Revista de Derecho Público: teoría y método* [en línea], 9, 171-220. Disponible en: https://doi.org/10.37417/RDP/vol_7_2024_2151
- San Martín Segura, D., 2023. Prevención algorítmica de la violencia de género: La discrecionalidad policial como decisión tecnológica en el contexto de VioGén. *Estudios Penales y Criminológicos* [en línea], 44(Número Especial), 1-35. Disponible en: <https://doi.org/10.15304/epc.44.9013>
- Solar Calvo, P., y Lacal Cuenca, P., 2021. Técnicas actuariales y valoración de peligrosidad. ¿Es este el camino? *Revista de Estudios Penitenciarios*, 263.
- Unión Europea, 2024. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial). *Diario Oficial de la Unión Europea*, 1689.