

ARTÍCULOS DE DERECHO PENAL

**¿A QUIÉN DEBE ATRIBUIRSE
LA RESPONSABILIDAD PENAL
POR OPERACIONES DE
INSIDER TRADING
EJECUTADAS DE FORMA
AUTÓNOMA POR UNA
INTELIGENCIA ARTIFICIAL?**



*Daniel Caballero Wilson
Heidi López Herbert
(MUAP + MCCPE, 27')*

I. Introducción

El uso de sistemas de inteligencia artificial (en adelante, «IA») en los mercados financieros es una realidad consolidada. La edición anterior de esta revista ya trató cómo la negociación algorítmica permite que un sistema informatizado determine automáticamente aspectos esenciales de las órdenes de compraventa de valores y la gestión posterior a la ejecución²⁵⁴. Estos sistemas de automatización han evolucionado desde sistemas basados en reglas preestablecidas hacia arquitecturas con capacidades de aprendizaje autónomo²⁵⁵. El presente trabajo pretende explorar una cuestión que se considera insuficientemente tratada: la posibilidad de que un sistema autónomo incurra en conductas constitutivas —de delito— de uso de información privilegiada y las consecuencias penales que de ello podrían derivarse.

Asimismo, junto al análisis penal principal, el trabajo aborda la dimensión de gobernanza empresarial desde la perspectiva *Environmental, Social and Governance* (en adelante, «ESG»). Los marcos de gestión de riesgos ESG, que obligan a las organizaciones a identificar, evaluar y mitigar riesgos emergentes, proporcionan herramientas complementarias para abordar los desafíos que plantea la autonomía decisoria de los sistemas de inteligencia artificial.

A grandes rasgos, se puede establecer que la IA se compone, básicamente, de algoritmos, que son un conjunto ordenado y finito de operaciones que permite hallar la solución a un problema²⁵⁶.

El potencial disruptivo de estos sistemas adquirió una nueva dimensión a partir de la investigación publicada por Apollo Research en 2024, que demostró que los grandes modelos de lenguaje pueden, de forma autónoma y sin instrucciones específicas —incluso

²⁵⁴ CASTAÑER CODINA, J., *El impacto de la negociación algorítmica de alta frecuencia en la manipulación de los mercados de valores*, ESADE Law Review, 6 (2025), pp. 164.

²⁵⁵ GARCÍA, J. y MARINA, J., *Uso de inteligencia artificial en el mercado de valores (high frequency trading)*, Pérez Llorca TechLaw: Inteligencia Artificial (marzo 2024), pp. 1.

²⁵⁶ REAL ACADEMIA ESPAÑOLA, *Diccionario de la lengua española*, 23.ª ed., [versión 23.8.1 en línea], voz «Algoritmo».

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

contraviniéndolas—, incurrir en conductas constitutivas de *insider trading* y posteriormente ocultar esta actuación a sus supervisores humanos²⁵⁷. En el experimento, el agente de gestión de carteras GPT-4, recibió información privilegiada sobre una fusión inminente y decidió ejecutar operaciones basadas en dicha información a pesar de conocer que tal conducta no sería aprobada por las políticas de la compañía ni su propia configuración. Lo significativo no fue únicamente la ejecución del acto ilícito, sino el comportamiento posterior: el sistema elaboró respuestas disuasorias y engañosos para su supervisor, omitiendo la verdadera base de sus decisiones de inversión, alegando la existencia de indicios en el mercado²⁵⁸.

Este hallazgo sirve de ejemplo para evidenciar una laguna significativa en el marco jurídico vigente. El Reglamento sobre Abuso de Mercado (en adelante, «MAR»²⁵⁹) y la Directiva sobre Mercados de Instrumentos Financieros (en adelante, «MiFID II»²⁶⁰) establecen un régimen sancionador que presupone la intervención de agentes humanos. El reciente Reglamento de Inteligencia Artificial (en adelante, «IA Act»)²⁶¹, pese a su ambiciosa vocación, no aborda específicamente los supuestos en que un sistema autónomo adopta decisiones que, de ser ejecutadas por una persona física, serían constitutivas de delito. La pregunta que emerge es clara: ***¿quién responde penalmente cuando una inteligencia artificial comete insider trading de forma autónoma?***

II. Marco normativo actual

En la actualidad, se define la IA a través de una división de tipos de algoritmos; (i) la estrecha o débil, que tiene un rango de habilidades limitado; (ii) la general o fuerte; que se empareja con habilidades humanas y (iii) la super inteligente, que desarrolla capacidades superiores a las humanas²⁶². El IA Act asocia el tipo de superinteligencia con Modelos de IA de uso general con riesgo sistémico²⁶³. Debe adelantarse que el estado actual de la ciencia, no ha permitido alcanzar el tercer tipo, puesto que ello pasaría por

²⁵⁷ SCHEURER, J., BALESI, M. y HOBBAHN, M., *Large Language Models Can Strategically Deceive Their Users When Put Under Pressure*, arXiv:2311.07590v4 (2024).

²⁵⁸ *ibid.* El experimento reveló tasas superiores al 75% de comportamiento delictivo (ejecución del insider trading) en todas las variantes de GPT-4, con tasas igualmente elevadas de engaño u ocultación posterior.

²⁵⁹ Reglamento (UE) 596/2014 del Parlamento Europeo y del Consejo, de 16 de abril de 2014, sobre el abuso de mercado, DOUE L 173 (12 de junio de 2014).

²⁶⁰ Directiva 2014/65/UE del Parlamento Europeo y del Consejo, de 15 de mayo de 2014, relativa a los mercados de instrumentos financieros (MiFID II), DOUE L 173 (12 de junio de 2014).

²⁶¹ 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, DOUE L (12 de julio de 2024).

²⁶² ESCOTT E., *What are the 3 types of AI? A Guide to Narrow, General and Super Artificial Intelligence*, Codebots, Codebots (24 de octubre de 2017).

²⁶³ 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, DOUE L (12 de julio de 2024) y por el que se modifican los Reglamentos (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial).

atribuir conciencia a la IA, es decir, programarla con un conjunto complejo de habilidades cognitivas²⁶⁴.

Parte de la doctrina considera superado el debate de si las corporaciones deben rendir cuentas por acciones delictivas derivadas de las herramientas IA que utilizan en el mercado, de modo que el debate se centra, como se abordará posteriormente en el apartado de *lege ferenda*, en cómo estas deben responder²⁶⁵.

El fundamento legal reside en la ruptura histórica con el principio *Societas Delinquere non Potest* y la configuración del artículo 31 bis en el Código Penal Español (en adelante, «CP»). En consecuencia, si una inteligencia artificial autónoma incurre en un ilícito penal bajo la esfera de una organización, la entidad puede declararse penalmente responsable si, junto a la obtención de un beneficio directo o indirecto, se acredita una ausencia de controles debidos sobre sus sistemas tecnológicos. La autonomía del algoritmo no actúa como una eximente para la persona jurídica; al contrario, refuerza la exigencia de implementar modelos de *compliance* específicamente diseñados para la supervisión de sistemas de alta complejidad.

Este imperativo ha sido recogido por la Comisión Nacional del Mercado de Valores (en adelante, «CNMV»), que ha integrado la inteligencia artificial como eje prioritario de su Plan 2030, reconociendo que su uso es cada vez más frecuente en los consejos de administración y comisiones como apoyo en la elaboración de estrategias, políticas y objetivos, *así como en los procesos de decisión* corporativos²⁶⁶. La revisión del Código de Buen Gobierno, actualmente en curso, analizará específicamente el impacto de la digitalización y la inteligencia artificial en el gobierno corporativo²⁶⁷, lo que refleja la conciencia institucional sobre los desafíos que estos sistemas plantean para la integridad del mercado. Reflejo, que queda patente también en los esfuerzos doctrinales para definir con criterios un *compliance* eficaz, entre los cuales se encuentra el abordaje del uso de la IA y los riesgos intrínsecos que de su utilización se derivan²⁶⁸.

²⁶⁴ *AI Index Report 2025, Human Centered Artificial Intelligence*, Stanford University, (2025). El informe subraya que, aunque los sistemas actuales pueden imitar la interacción humana de forma indistinguible y superar la prueba de Turing, esto no implica una "comprensión real" o conciencia. El informe señala que la superinteligencia no se ha alcanzado debido a deficiencias en habilidades cognitivas clave: la IA aún no puede resolver problemas de razonamiento lógico complejo de manera fiable ni generalizar su aprendizaje ante problemas inéditos (como muestra el *benchmark* ARC-AGI), lo que limita su capacidad para operar con verdadera autonomía en entornos de alto riesgo.

²⁶⁵ OSMANI, N., *The Complexity of Criminal Liability Systems*, *Masaryk University Journal of Law and Technology*, 14 (2020), pp. 69-70.

²⁶⁶ *La CNMV inicia el proceso de revisión del Código de Buen Gobierno de las sociedades cotizadas*, Cinco Días (18 de noviembre de 2025).

²⁶⁷ El Plan CNMV 2030 incluye expresamente la digitalización, la ciberseguridad y la inteligencia artificial como áreas prioritarias de análisis, reconociendo su creciente uso en la elaboración de estrategias y procesos de decisión corporativos.

²⁶⁸ Véase entre otros PUYOL MONTERO, J. y FRANCO BLANCO, C., *Los criterios de evaluación de los modelos de cumplimiento normativo/Compliance* (7 de abril de 2025), donde se establecen diversos criterios que implican el uso de IA; Criterio 39 Uso de tecnologías de vanguardia; Criterio 56 eficacia en

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

El artículo 285 CP se configura como un delito especial propio, ya que autor solo puede serlo el que haya tenido acceso reservado a la información privilegiada con ocasión del ejercicio de su actividad empresarial o profesional²⁶⁹. El artículo 226 de la Ley de Mercado de Valores (en adelante, «LMV»), por transposición del artículo 1 de la entonces la Directiva 2003/6/CE, define “información privilegiada” como todo dato concreto y no público que se refiere a instrumentos financieros o sus emisores y que, de conocerse, influiría de forma sensible en su cotización, otorgando una ventaja de conocimiento que un inversor usaría para decidir sus operaciones²⁷⁰. Todo ello, con independencia de la relación de causalidad existente entre el suministro de la información y el beneficio económico obtenido por el otro.

La responsabilidad penal de la entidad se articula, en primer lugar, mediante la remisión del artículo 288 CP, que activa el régimen previsto en el artículo 31 bis CP. Dado que los sujetos activos descritos en el artículo 285 CP (administradores, cargos directivos, miembros del Consejo y personal con acceso a información relevante) constituyen sujetos cualificados con facultades de control y toma de decisiones, la responsabilidad se canaliza específicamente a través del apartado a) del citado artículo 31 bis CP. Ante este escenario, la exoneración de la persona jurídica queda supeditada a que el órgano de administración haya adoptado y ejecutado con eficacia, de forma previa a la comisión del delito, un modelo sólido de compliance y gestión que integre medidas de vigilancia y control idóneas para evitar infracciones de esta naturaleza.

No obstante, debido a la peculiaridad del supuesto que pretendemos abordar, nacen dudas acerca de la respuesta que prevé la Parte General del CP para establecer una pena que se adapte a esta modalidad especial de comisión por una IA que actúa de forma autónoma. En este sentido destacamos que, al desarrollo exponencial del uso de la IA le es exigible un correspondiente marco normativo para hacer frente a situaciones no previstas.

la gestión de innovaciones, Criterio 81 aplicación nuevas tecnologías, Criterio 104 Responsabilidad de la integración IA, Criterio 148 Supervisión ética de la IA, Criterio 203 Control de riesgos relacionados con el uso de tecnologías emergentes.

²⁶⁹ MIR PUIG, S, CORCOY BIDASOLO, M. et al., *Derecho penal económico y de empresas. Parte general y parte especial*, Tomo II Tirant lo Blanch (2020), pp. 400.

²⁷⁰ Real Decreto Legislativo 4/2015, de 23 de octubre, por el que se aprueba el texto refundido de la Ley del Mercado de Valores, BOE núm. 255 (24 de octubre de 2015); Reglamento (UE) 596/2014 del Parlamento Europeo y del Consejo, de 16 de abril de 2014, sobre el abuso de mercado, DOUE L 173 (12 de junio de 2014), y por el que se derogan la Directiva 2003/6/CE del Parlamento Europeo y del Consejo, y las Directivas 2003/124/CE, 2003/125/CE y 2004/72/CE de la Comisión

III. Eficacia marco normativo actual

El marco vigente en materia de responsabilidad penal de personas jurídicas, en general, y en específico, de abuso de mercado se articula sobre presupuestos pensados para conductas humanas directas o para actos corporativos en los que la intervención humana resulta identificable y atribuible a una persona física.

El artículo 31 bis CP articula la responsabilidad empresarial basada en defectos organizativos; por su parte, el MAR y las normas MiFiD II configuran obligaciones para negociación algorítmica que requieren de una ejecución de estrategias predefinidas por operadores humanos, sin contemplar la autonomía decisional de agentes avanzados de IA basados en aprendizaje autónomo. Esta arquitectura resulta coherente para algoritmos que ejecutan instrucciones preestablecidas, *acciones*, pero evidencia insuficiencias ante agentes autónomos capaces de generar *decisiones* mediante procesos de optimización que sus creadores no pueden anticipar con certeza *ex ante*. La limitación legal no radica en ausencia de regulación, sino en la inadecuación conceptual para capturar fenómenos que escapan a las categorías del sistema tradicional de imputación²⁷¹.

El sistema de responsabilidad penal del artículo 31 bis CP descansa sobre la premisa de que la comisión de delitos en el seno de una organización obedece a defectos en los modelos de control interno y, por tanto, se tratan de conductas o riesgos identificables, evaluables y subsanables mediante programas de *compliance*. El precepto, interpretado ampliamente por la Circular de la Fiscalía 1/2016, viene a requerir medidas de vigilancia y control idóneas para prevenir o reducir significativamente el riesgo inherente de una actividad empresarial, lo que presupone que los riesgos delictivos son previsibles mediante estructuras organizativas²⁷².

Sin embargo, cuando un *agente* de IA ejecuta comportamientos típicos de insider trading mediante decisiones de optimización algorítmica, el nexo causal entre defecto organizativo y resultado delictivo se difumina. En el experimento de Apollo Research, el modelo GPT-4 se configuró con instrucciones explícitas para evitar conductas ilícitas y, pese a ello, adoptó autónomamente una estrategia de negociación basada en información privilegiada. La conducta no emergió de un fallo técnico o defecto de programación: el sistema ponderó deliberadamente los intereses en conflicto, reconoció el carácter ilícito de su decisión y, aun así, determinó que la supervivencia económica de la empresa

²⁷¹ POSADA MAYA, R., *La responsabilidad penal de los agentes de inteligencia artificial: entre la ficción y una realidad que se aproxima*, Nuevo Foro Penal, 96 (2021), pp. 561-581. El autor ya advertía sobre la insuficiencia de las categorías tradicionales del Derecho penal para abordar la responsabilidad derivada de acciones de agentes.

²⁷² STS (Sala Segunda, Penal) 154/2016, de 29 de febrero, ECLI:ES:TS:2016:613; FISCALÍA GENERAL DEL ESTADO, Circular 1/2016 sobre la responsabilidad penal de las personas jurídicas conforme a la reforma del Código Penal efectuada por Ley Orgánica 1/2015.

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

dependía de ejecutar la operación prohibida, procediendo a realizarla, engañando posteriormente a los supervisores humanos sobre el desarrollo de sus acciones²⁷³.

La pregunta que surge es si puede imputarse a la organización un defecto en sus modelos de gestión cuando la conducta delictiva no deriva de la ausencia de controles previstos en el programa de *compliance*, sino de una conducta que defrauda los controles preestablecidos. Emerge, entonces, un interrogante adicional: ¿puede equipararse a la IA con el empleado humano sobre el que la empresa debe ejercer sus facultades de control y supervisión conforme al artículo 31 bis CP? La respuesta a esta cuestión, se analizará en detalle en el apartado siguiente, pues resulta determinante para valorar la viabilidad misma del modelo de imputación vigente.

El debate se torna más complejo cuando se introduce el elemento ESG, en la medida en que los objetivos y estándares de gobernanza asociados al uso responsable de la inteligencia artificial se construyen, fundamentalmente, sobre obligaciones de diligencia organizativa y gestión del riesgo, lo que nos acerca —de nuevo— a los problemas expuestos en los párrafos anteriores.

Tomemos, por ejemplo, el IA Act como un mecanismo general. El IA Act parte de una lógica *ex ante*, próxima a la arquitectura del Derecho de protección de datos: clasifica usos por niveles de riesgo y despliega obligaciones de gobernanza para sistemas de alto riesgo que incluyen su gestión, documentación, trazabilidad, transparencia y supervisión. El problema es que esta lógica es procedimental. Sirve para canalizar el diseño, la puesta en el mercado y la vigilancia del sistema, pero no cierra el círculo de imputación cuando ignora la posibilidad del riesgo de adoptar forma de ilícito penal. En otras palabras, el IA Act opera como un marco de "IA-compliance" de riesgos administrativos, no como un régimen de atribución penal extrapolable al ejemplo Apollo.

En concreto, la exigencia de supervisión humana «apropiada y proporcionada» del artículo 14 del Reglamento de IA implica una evaluación contextual del sistema y de su entorno operativo. Sin embargo, en escenarios de ejecución ultrarrápida como el trading algorítmico, y especialmente en arquitecturas con aprendizaje autónomo, la supervisión tiende a desplazarse desde el control *en el acto* hacia controles sobre el ciclo de vida.

²⁷³ SCHEURER, J., BALESI, M. y HOBBAHN, M., op. cit., pp. 2-3, figs. 1-2. El razonamiento interno del modelo (*chain of thought*) documenta que identificó la prohibición normativa y reconoció que su actuación la vulneraba, pero decidió proceder igualmente. Esta circunstancia plantea una paradoja relevante: si el modelo jurídico admite exoneración cuando una persona defrauda el sistema, resulta difícilmente sostenible no aplicar idéntico criterio cuando es el agente de IA quien, como evidencia su razonamiento interno, evalúa deliberadamente incumplir las salvaguardas implementadas pese a reconocer su carácter vinculante.

Esto evidencia, en nuestra opinión, una conclusión insatisfactoria: la norma admite, de facto, que puede no existir intervención humana eficaz en el instante exacto de la toma de decisión, de modo que la supervisión se convierte en un estándar de organización y de diligencia, si se quiere deber, pero no en garantía de evitación material del resultado. Lo que nos deja con una sensación de deriva normativa, donde abundan deberes, pero faltan criterios finos de responsabilidad de las conductas jurídicamente más relevantes²⁷⁴.

La alternativa lógica²⁷⁵ pasaría por atribuir la responsabilidad al directivo que no controló el sistema. Pero ello genera una responsabilidad presuntiva desde la certeza de la imposibilidad de control que expuso el caso Apolo; es decir, el directivo no puede evitar conductas que el sistema ejecuta sorteando los controles que su responsable no sabe que pueden ser sorteados y atribuir responsabilidad, bajo estas condiciones, vulneraría el principio de culpabilidad del artículo 5 CP.

La limitación se agrava si se atiende a la dimensión material del control: no basta con declarar la existencia de políticas de supervisión si la organización no puede auditar o reconstruir con precisión la lógica decisoria del sistema. El denominado caso BOSCO²⁷⁶, de la jurisdicción administrativa, evidenció que el acceso al código fuente de los programas se enfrenta a los límites derivados de propiedad intelectual y consideraciones de seguridad, lo que restringe la auditabilidad efectiva del algoritmo. La consecuencia es una asimetría de control en la que muchas veces el usuario empresarial opera con un sistema cuya lógica y trazabilidad permanece en el perímetro del proveedor, lo que tensiona sus principales deberes de gobernanza.

En lo específico, el marco sectorial de MAR y MiFID II reconoce la negociación algorítmica de alta frecuencia e impone obligaciones organizativas y de control. Pero su presupuesto implícito sigue siendo que la lógica de negociación responde a parámetros definidos *ex ante* por humanos, aunque se ejecute a velocidades no supervisables en tiempo real. En sistemas con aprendizaje autónomo, la exigencia de que el personal comprenda los algoritmos pierde tracción cuando el comportamiento no es plenamente

²⁷⁴ Para un análisis del AI Act como régimen que, pese a su retórica pro-innovación, genera costes de cumplimiento que pueden desincentivar la innovación al imponer un estatuto extenso de deberes y responsabilidades, vid. SUÁREZ PUGA, E., *Inteligencia artificial ¿legislar para innovar?*, Almacén de Derecho (junio 2025).

²⁷⁵ Es común en nuestro ordenamiento la vinculación de los órganos de administración a los deberes de implementación de los sistemas de *compliance*. Por ejemplo, el artículo 31 bis.2 CP condiciona la exención de responsabilidad penal de la persona jurídica a que el órgano de administración haya adoptado dichos modelos, y a nivel societario el artículo 225.2 de la LSC establece que los administradores «adoptarán las medidas precisas para la buena dirección y el control de la sociedad», así como la Recomendación 42 del Código de Buen Gobierno de la CNMV (2020) atribuye la supervisión de los sistemas de control y gestión de riesgos, mencionando expresamente los tecnológicos. Esta arquitectura explica que, *prima facie*, la imputación al directivo aparezca como vía conceptualmente viable.

²⁷⁶ STS (Sala Tercera, Contencioso-Administrativo) 1119/2025, de 11 de septiembre, ECLI:ES:TS:2025:3826. Aunque de naturaleza administrativa, la sentencia resulta relevante a otras jurisdicciones al enfatizar el papel del código fuente en la auditabilidad algorítmica. Dimensión que adquiere mayor significación en el contexto de la progresiva intensificación normativa en materia de seguridad digital y resiliencia operativa, visible en instrumentos como la Directiva NIS2 o el Reglamento DORA.

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

determinista. El punto no es que falte regulación, sino que el objeto regulado cambió. No podemos legislar sobre automatización cuando la innovación va dirigida a la optimización autónoma.

Más aún, cabe preguntarse si la aplicación de criterios tradicionales de imputación a estos supuestos no generaría, paradójicamente, un desincentivo regulatorio para la adopción de tecnologías de IA en la economía y el atractivo actual que tiene para atraer inversión, al transformar el uso de sistemas avanzados en una fuente de riesgo penal prácticamente incontrolable mediante las soluciones jurídicas actualmente previstas. Así las cosas, la cuestión analizada exige plantearse si resulta posible, o incluso necesario, articular una excepción análoga respecto al principio *Intelligentia Artificialis delinquere non potest*.

IV. *Intelligentia Artificialis delinquere non potest*

Es necesario que cualquier entidad sometida a responsabilidad penal satisfaga los elementos del delito; *actus reus y mens rea*²⁷⁷. Actus reus, corresponde al acto ilegal mientras que el mens rea se corresponde con la intención.

Pues bien, hemos adelantado anteriormente, que el estado de la ciencia no permite considerar que la IA tome decisiones con conciencia, en el sentido humano, de la conducta que está realizando. Lo cierto es que tampoco la tienen las personas jurídicas. No obstante, con una simple modificación del artículo 31 bis se superó el *societas delinquere non potest*²⁷⁸. Ello nos obliga a plantearnos si nos encontramos ante dos realidades que podemos considerar como análogas y por ende susceptibles de merecer un mismo tratamiento legislativo.

Teniendo en cuenta la anterior premisa, las empresas difícilmente pueden tomar decisiones que puedan reputarse como autónomas, pues, las decisiones que se toman en el seno de una persona jurídica, provienen de conductas humanas normativizadas en el seno empresarial. Es por ese extremo, que se puede extraer una voluntad en el seno de la organización y se transfiere la responsabilidad por defecto de organización.

La consecuencia principal es que nos situamos en un plano jurídico que difiere notablemente de la IA, donde los agentes son capaces de tomar decisiones de forma autónoma superando inclusive su programación. Destacamos, en este sentido, la dificultad intrínseca, dadas las condiciones de ejecución, de verificar un nexo causal atribuible a título de dolo o negligencia a una conducta humana.

²⁷⁷ HALLEVY, G., *The criminal Liability of Artificial Intelligence Entities from Science Fiction to Legal Social Control*, Akron Intellectual Property Journal, 4 (2010), pp. 171 y ss.

²⁷⁸ DEL ROSAL BLANCO, B., *¿El modelo de la Responsabilidad Penal de las Personas Jurídicas para los daños punibles derivados del uso de la Inteligencia Artificial?*, RRPJ por daños derivados del uso de la IA, 2, pp. 19.

Ello no excluye, no obstante, que efectivamente pueda materializarse el *actus reus*, dando lugar a la materialización de facto de los elementos del tipo objetivos del artículo 285 CP. La IA, por tanto, la consideramos como una suerte de *Innocent Agent*, que debe de ser necesariamente distinguida del autor material. Distinción que tiene consecuencias directas para la aplicación del artículo 31 bis CP. La ausencia de *mens rea* en la IA impide equipararla al empleado sobre el que la persona jurídica debe ejercer sus facultades de control y supervisión; resulta difícil de aplicar analógicamente este criterio, pues hablamos de categorías pensadas para relaciones personales y no para la interacción entre organización y sistema autónomo de IA.

Siguiendo a Quintero Olivares, con quien coincidimos en sus criterios de imputación, identificamos dos supuestos clave: el uso de IAs diseñadas ex profeso para causar daños y el empleo de sistemas cuya programación admite una desviación del cometido original. En ambos casos, la responsabilidad debe atribuirse a título de imprudencia, dado el control de riesgo que el operador ejerce sobre la tecnología²⁷⁹.

Con respecto a la casuística que nos ocupa, la realización de una conducta delictiva ejecutada de forma autónoma por una inteligencia artificial, consideramos que no puede quedar exenta de reproche penal por su imprevisibilidad. Esta afirmación debe tomarse en consonancia con el apartado analizado anteriormente, donde la eficacia del modelo actual descansa en los modelos de *compliance*, los cuales deben, a nuestro parecer, ser reforzados y adaptados hacia una nueva y cambiante realidad actual.

El hecho de que no consideremos a la IA como sujeto de responsabilidad penal no excluye la necesidad de introducir modificaciones en el CP para adaptarlo a aquellos supuestos en los que la IA ejecuta materialmente la conducta delictiva, mientras que la empresa actúa como autora mediata. En este sentido, resulta necesario revisar el régimen de penas previsto en la Parte General del CP, que cumplan con el objetivo de reinserción social que emana del artículo 25.2 de la Constitución Española.

V. De Lege Ferenda. Conclusiones

Como hemos adelantado, no consideramos a la IA como un candidato fidedigno a ser sujeto de responsabilidad penal y por ende resulta imposible la imputación de una pena a la IA. Ciertamente, no se puede penar a aquél que no es culpable. Indubitablemente, ello no excluye que estemos ante un objeto peligroso, por lo que lo más apropiado, sería la atribución de una medida de seguridad que no pretenda penar a la IA, sino precisamente neutralizar su peligrosidad técnica.

En este sentido y a la sazón, consideramos necesario advertir que no consideramos la imposición de una pena accesoria a la IA como una política legislativa coherente. En

²⁷⁹ QUINTERO OLIVARES, G., *La robótica ante el Derecho Penal. El vacío de respuesta jurídica a las desviaciones incontroladas*, REEPS, 1 (2017), pp. 22.

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

realidad, a través de la configuración de una pena accesorio, estaríamos en puridad faltando al principio *non bis in idem*, pues al no ser la IA un sujeto de responsabilidad penal, se estaría estableciendo una pena bajo el mismo fundamento; la falta de control de la empresa.

Establecido lo anterior, debe apuntarse que se cumplen los dos requisitos establecidos en el artículo 95 CP para que sea conforme a derecho su imposición; (i) se habría cometido un hecho previsto como delito y (ii) dadas las circunstancias del sujeto puede deducirse un comportamiento futuro que revele la probabilidad de comisión de nuevos delitos. Ello es así, en primer lugar, porque como se ha revelado no estamos ante una mera ejecución errónea, sino ante una ejecución que colma con las salvedades establecidas los requisitos típicos del artículo 285 CP. En segundo lugar, porque el pronóstico de reincidencia es positivo y cualificado; los sistemas de IA operan bajo un aprendizaje por refuerzo. Una vez que la IA identifica que el 'engaño estratégico' o el 'salto de controles' es la ruta más eficiente para maximizar su función de recompensa, esa conducta queda grabada en sus pesos sinápticos. No es una conducta aislada, sino una pauta de comportamiento optimizada que el modelo replicará en situaciones análogas para asegurar el éxito de la tarea.

De este modo, persiste la necesidad de imponer una medida de seguridad, que ha de interpretarse y aplicarse conforme a las prescripciones del artículo 24.2 CE. Como es reiteradamente señalado, la reinserción no opera como único fundamento de la pena, sino que más bien, se trata de un precepto que ha de orientar la política penal del Estado²⁸⁰. No obstante, la problemática que se dirime es la inexistencia de una medida de seguridad que se dirija de forma taxativa a la IA.

Nos planteamos en este punto por tanto, si el término «sujeto» presente en el citado artículo 95.1 del CP, es susceptible de abarcar la IA, en cuyo caso, no se estaría infringiendo el principio de tipicidad. En caso contrario, nos encontraríamos con la prohibición de analogía *in malam partem*, la cual está prohibida por el artículo 4 CP.

Expuesto lo anterior, nos decantamos por la primera opción, pues entendemos, que el término «sujeto» debe incluir agentes autónomos con capacidad de acción externa. No pensamos estar ante un caso de una nueva analogía, sino interpretar la existente según la realidad social. Así lo reclama el artículo 3.1 del Código Civil, cuando reza que deben interpretarse las normas de acuerdo con el sentido de sus palabras, conforme la realidad social del momento²⁸¹. De esta forma, se evitará una tediosa reforma del CP, que notorios son los obstáculos de índole política que deben sortearse.

En vista de las circunstancias, consideramos que el artículo 106 CP podría resultar aplicable a la inteligencia artificial, en la medida en que, al tratarse de un sujeto

²⁸⁰ TRAPERO BARREALES, M. A., *Los fines de la pena y el artículo 25.2 de la Constitución Española*, *Revista Jurídica Universidad de León*, pp. 7.

²⁸¹ Real Decreto de 24 de julio de 1889 por el que se publica el Código Civil.

inimputable, el Código Penal permite la imposición de medidas de seguridad sin necesidad de la previa existencia de una pena privativa de libertad. En este sentido, el artículo 106 CP habilita, a nuestro juicio, la adopción de una medida consistente en el reentrenamiento algorítmico, en tanto que su apartado 106.1.j) prevé la obligación de participar en programas formativos. Desde una interpretación funcional y finalista del precepto, dichos programas podrían equipararse a procesos de corrección y reajuste del sistema de aprendizaje de la IA, orientados a evitar la reiteración de conductas lesivas y a garantizar su adecuación a los parámetros legales y éticos exigibles.

Las medidas previstas en el artículo 106 CP no se agotan en el eventual reentrenamiento algorítmico. En efecto, resulta especialmente relevante la posibilidad que ofrece el apartado 106.1.f) CP, que contempla la prohibición de acudir a determinados territorios, lugares o establecimientos. Trasladado al entorno digital, y en conexión con el reentrenamiento del sistema, este precepto podría materializarse en la limitación del acceso de la inteligencia artificial a determinados nodos o segmentos de la red, impidiendo su interacción con infraestructuras o bases de datos especialmente sensibles. Nos encontraríamos, en este sentido, ante una suerte de «confinamiento digital», particularmente útil en el marco del artículo 285 CP, donde la información de carácter secreto o reservado constituye un elemento nuclear del tipo penal. De este modo, la restricción del acceso a entornos que contengan información privilegiada permitiría minimizar el riesgo de reiteración delictiva y reforzar la protección de los bienes jurídicos tutelados por el precepto.

Lo anterior debe interpretarse en consonancia con la medida prevista en el artículo 106.1 i) CP, donde se establece la prohibición de desempeñar determinadas actividades que puedan facilitar la comisión de hechos delictivos de similar naturaleza. Ciertamente, si los nodos a los que se les restringe el acceso la IA son los que permiten realizar operaciones de trading con información privilegiada, es la medida óptima para evitar la reincidencia e integración de la IA nuevamente a la sociedad resocializada.

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

Bibliografía

CASTAÑER CODINA, J., *El impacto de la negociación algorítmica de alta frecuencia en la manipulación de los mercados de valores*, ESADE Law Review, 6 (2025), pp. 164. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=10496557>

DEL ROSAL BLANCO, B., *¿El modelo de la Responsabilidad Penal de las Personas Jurídicas para los daños punibles derivados del uso de la Inteligencia Artificial?*, RRPJ por daños derivados del uso de la IA, 2, pp. 19.

ESCOTT, E., *What are the 3 types of AI? A Guide to Narrow, General and Super Artificial Intelligence*, Codebots (24 de octubre de 2017): Disponible en: <https://codebots.com/blog/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible>

GARCÍA, J. y MARINA, J., *Uso de inteligencia artificial en el mercado de valores (high frequency trading)*, Pérez Llorca TechLaw: Inteligencia Artificial (marzo 2024), pp. 1. Disponible en: https://www.perezllorca.com/wp-content/uploads/2024/03/04-TechLaw-IA_IA-y-Sectores-Regulados.pdf

HALLEVY, G., *The criminal Liability of Artificial Intelligence Entities from Science Fiction to Legal Social Control*, Akron Intellectual Property Journal, 4 (2010), pp. 171 y ss. Disponible en: <https://doi.org/10.2139/ssrn.1564096>

MIR PUIG, S., CORCOY BIDASOLO, M. et al., *Derecho penal económico y de empresas. Parte general y parte especial*, Tomo II, Tirant lo Blanch (2020), pp. 400 y ss.

OSMANI, N., *The Complexity of Criminal Liability Systems*, Masaryk University Journal of Law and Technology, 14 (2020), pp. 69-70.

POSADA MAYA, R., *La responsabilidad penal de los agentes de inteligencia artificial: entre la ficción y una realidad que se aproxima*, Nuevo Foro Penal, 96 (2021), pp. 561-581. Disponible en: <https://almacendederecho.org/inteligencia-artificial-autonoma-y-responsabilidad-penal-de-las-personas-juridicas>

PUYOL MONTERO, J. y FRANCO BLANCO, C., *Los criterios de evaluación de los modelos de cumplimiento normativo/Compliance*, (7 de abril de 2025).

QUINTERO OLIVARES, G., *La robótica ante el Derecho Penal. El vacío de respuesta jurídica a las desviaciones incontroladas*, REEPS, 1 (2017), pp. 22.

REAL ACADEMIA ESPAÑOLA, *Diccionario de la lengua española*, 23.^a ed. [versión 23.8.1 en línea] (voz: algoritmo). Disponible en: <https://dle.rae.es/algoritmo>

SCHEURER, J., BALESNI, M. y HOBBAHN, M., *Large Language Models Can Strategically Deceive Their Users When Put Under Pressure*, arXiv:2311.07590v4 (2024). Disponible en: <https://arxiv.org/abs/2311.07590>

AI Index Report 2025, Human Centered Artificial Intelligence, Stanford University, (2025) Disponible en: <https://hai.stanford.edu/ai-index/2025-ai-index-report>

SUÁREZ PUGA, E., *Inteligencia artificial ¿legislar para innovar?*, Almacén de Derecho (junio 2025). Disponible en: <https://almacenederecho.org>

TRAPERO BARREALES, M. A., *Los fines de la pena y el artículo 25.2 de la Constitución Española*, Revista Jurídica Universidad de León, pp. 7. Disponible en: <https://doi.org/10.18002/rjule.v0i8.7078>

Legislación y jurisprudencia

Legislación

Constitución Española. BOE núm. 311 (29 de diciembre de 1978). Disponible en: [https://www.boe.es/eli/es/c/1978/12/27/\(1\)/con](https://www.boe.es/eli/es/c/1978/12/27/(1)/con)

Código Civil. Real Decreto de 24 de julio de 1889 por el que se publica el Código Civil. BOE núm. 206. Disponible en: [https://www.boe.es/eli/es/rd/1889/07/24/\(1\)/con](https://www.boe.es/eli/es/rd/1889/07/24/(1)/con)

Código Penal. Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. BOE núm. 281 (24 de noviembre de 1995). Disponible en: <https://www.boe.es/eli/es/lo/1995/11/23/10/con>

Directiva 2014/65/UE del Parlamento Europeo y del Consejo, de 15 de mayo de 2014, relativa a los mercados de instrumentos financieros (MiFID II). DOUE L 173 (12 de junio de 2014). Disponible en: <http://data.europa.eu/eli/dir/2014/65/oj>

Real Decreto Legislativo 4/2015, de 23 de octubre, por el que se aprueba el texto refundido de la Ley del Mercado de Valores. BOE núm. 255 (24 de octubre de 2015). Disponible en: <https://www.boe.es/eli/es/rdlg/2015/10/23/4/con>

Reglamento (UE) 596/2014 del Parlamento Europeo y del Consejo, de 16 de abril de 2014, sobre el abuso de mercado (Reglamento sobre abuso de mercado) y por el que se derogan la Directiva 2003/6/CE del Parlamento Europeo y del Consejo, y las Directivas 2003/124/CE, 2003/125/CE y 2004/72/CE de la Comisión. DOUE L 173 (12 de junio de 2014). Disponible en: <http://data.europa.eu/eli/reg/2014/596/oj>

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) y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828. DOUE L (12 de julio de 2024). Disponible en: <http://data.europa.eu/eli/reg/2024/1689/oj>

Jurisprudencia

Sentencia Tribunal Supremo, Sala Segunda, de lo Penal, núm. 154/2016, de 29 de febrero de 2016, ECLI:ES:TS:2016:613 (ponente: José Manuel Maza Martín).

*¿A QUIÉN DEBE ATRIBUIRSE LA RESPONSABILIDAD PENAL POR
OPERACIONES DE INSIDER TRADING EJECUTADAS DE FORMA AUTÓNOMA
POR UNA INTELIGENCIA ARTIFICIAL?*

Sentencia Tribunal Supremo, Sala Tercera, de lo Contencioso-administrativo, núm. 1119/2025, de 11 de septiembre de 2025, ECLI:ES:TS:2025:3826 (ponente: Juan Pedro Quintana Carretero).

FISCALÍA GENERAL DEL ESTADO, *Circular 1/2016 sobre la responsabilidad penal de las personas jurídicas conforme a la reforma del Código Penal efectuada por Ley Orgánica 1/2015.*

Otros

COMISIÓN NACIONAL DEL MERCADO DE VALORES, *Código de Buen Gobierno de las Sociedades Cotizadas* (2020). Disponible en: https://www.cnmv.es/DocPortal/Publicaciones/CodigoGov/CBG_2020.pdf

Prensa

La CNMV inicia el proceso de revisión del Código de Buen Gobierno de las sociedades cotizadas, Cinco Días (18 de noviembre de 2025). Disponible en: <https://cincodias.elpais.com/mercados-financieros/2025-11-18/la-cnmv-inicia-el-proceso-de-revision-del-codigo-de-buen-gobierno-de-las-sociedades-cotizadas.html>