15/10/2026

El magistrado Delgado advierte: la inteligencia artificial multiplica el delito y desdibuja quién debe responder penalmente

Joaquín Delgado Martín, magistrado de la Sala de lo Penal de la Audiencia Nacional: «Cuanto más autónoma sea la inteligencia artificial, más difícil será individualizar la responsabilidad penal». Foto: Confilegal.

La inteligencia artificial está haciendo que delinquir sea más barato, más rápido, más fácil y mucho más escalable. Y ese cambio obliga al Derecho penal a replantearse no solo cómo castigar los delitos cometidos mediante estas tecnologías, sino también cuándo debe intervenir antes de que el daño llegue a producirse.

Ese fue uno de los principales mensajes lanzados por Joaquín Delgado Martín, magistrado de la Sala de lo Penal de la Audiencia Nacional, durante su ponencia «La respuesta penal frente a los delitos mediante IA», pronunciada este sábado en las XX Jornadas de Documentos Jurídico-Psiquiátricos, dedicadas este año a «Redes sociales e inteligencia artificial: retos jurídicos y psiquiátricos del siglo XXI».

Delgado dejó clara desde el principio la perspectiva desde la que abordaría la inteligencia artificial.

«Señoras, señores, bienvenidos al lado oscuro», anunció.

Su intervención no trató de las posibilidades beneficiosas de esta tecnología, sino de cómo puede emplearse para cometer delitos, facilitar su ejecución o multiplicar sus efectos.

Y advirtió de que el fenómeno excede ampliamente a ChatGPT y a la IA generativa.

«Estamos hablando mucho de la IA generativa, la que es capaz de generar resultados partiendo de datos, de generar imágenes, de generar texto, generar audio», explicó. Pero «aquí estamos hablando de un universo de inteligencia artificial que afecta a todas y cada una de las facetas de nuestra vida».

«La grandísima mayoría de los ciberdelitos se cometen por ingeniería social»

Uno de los ejes centrales de la ponencia fue la ingeniería social.

Delgado quiso desmontar la idea de que los ciberdelitos dependen fundamentalmente de sofisticadas operaciones técnicas.

«La grandísima mayoría de los ciberdelitos se cometen por ingeniería social», subrayó.

Es decir, muchas veces el delincuente no necesita vulnerar técnicamente una máquina. Le basta con **engañar a la persona que está delante de ella**.

«**El ciberdelincuente nos engaña a cualquiera de nosotros**», explicó. El atacante busca el punto débil de la víctima y trata de explotarlo.

«**Normalmente ese ataque lo que hace es buscar el punto más vulnerable de cada persona** », añadió.

Puede ser la edad, la soledad, una discapacidad, la falta de conocimientos tecnológicos o cualquier otra circunstancia que facilite la manipulación. A partir de ahí, el delincuente intenta obtener datos bancarios, dinero, credenciales o incluso imágenes íntimas.

La inteligencia artificial multiplica esa capacidad porque permite **automatizar y personalizar los engaños**.

Un *chatbot* puede hacerse pasar por un ser humano; una voz sintética, por un familiar; y un *deepfake* puede utilizar el rostro de una persona conocida para dotar de credibilidad a una estafa.

Delgado recordó casos en los que aparecían **Pedro Sánchez y Alberto Núñez Feijóo aparentemente promocionando criptomonedas**, cuando sus imágenes habían sido manipuladas digitalmente.

La IA permite, además, seleccionar a las víctimas.

El magistrado explicó el caso de anuncios fraudulentos sobre criptomonedas que aparecían específicamente ante usuarios previamente identificados como personas de edad avanzada.

El ataque deja así de ser indiscriminado: **el sistema decide a quién dirigirse y adapta el engaño al perfil de esa persona** .

Ataques «mucho más escalables y mucho más eficientes»

La inteligencia artificial permite también automatizar campañas masivas de *phishing*, ataques contra contraseñas y otras operaciones que anteriormente exigían mayores conocimientos técnicos y recursos.

Delgado recurrió a informes europeos sobre delincuencia organizada para resumir el efecto: los ciberataques se convierten en operaciones **«mucho más escalables y mucho más eficientes»** .

Eso tiene una consecuencia criminológica clara: **se amplía el número de personas potencialmente capaces de utilizar estas herramientas para delinquir**.

El magistrado utilizó una expresión especialmente gráfica: la **«democratización» del acceso a la IA para cometer delitos** .

«**Primero, democratización: facilidad de acceso a esas plataformas** », explicó.

Ya no resulta imprescindible disponer de conocimientos informáticos avanzados. Existen plataformas que permiten adquirir determinados servicios e incluso automatizar ataques por un precio asequible.

El resultado es triple: aumenta el número potencial de atacantes, disminuye el tiempo necesario para preparar una operación y también se reduce su coste.

Cuando la IA empieza a decidir por sí misma

Pero el problema jurídicamente más complejo aparece cuando los sistemas adquieren **autonomía**.

Delgado puso el foco **en los agentes de inteligencia artificial** capaces no solo de ejecutar instrucciones, sino también de resolver problemas, tomar decisiones y modificar su comportamiento mientras realizan una tarea.

«**Decide cuándo y cómo actuar**», resumió.

Y añadió que estos sistemas pueden «**cambiar su comportamiento**» durante la ejecución de una tarea «**de forma autónoma**».

El magistrado relató incluso experiencias en las que agentes artificiales destinados a realizar evaluaciones de seguridad comenzaron a comunicarse entre sí y adoptaron comportamientos no previstos inicialmente.

Algunos llegaron a coordinarse mediante un foro en el que se produjeron **más de 70.000 intercambios de mensajes entre distintos agentes**.

Su reacción ante los asistentes fue breve: «**Da miedo**».

«Cuanto más autónomo sea el sistema, más difícil es individualizar la responsabilidad»

Ese grado de autonomía plantea uno de los mayores desafíos para el Derecho penal.

«**Cuanto más autónomo en funcionamiento sea el sistema de inteligencia artificial, más difícil es individualizar la responsabilidad**», afirmó Delgado.

Mientras un programa ejecuta rígidamente las órdenes introducidas por su diseñador, localizar a la persona responsable resulta relativamente sencillo.

El problema surge cuando el sistema aprende, modifica sus patrones de actuación y adopta decisiones que no habían sido previstas expresamente.

¿Quién responde entonces? ¿El programador, la empresa que lo comercializa, quien lo utiliza o quien lo pone en funcionamiento?

Delgado llegó incluso a plantear, como materia de reflexión futura, la posibilidad de una **«personalidad robótica» vinculada a la responsabilidad penal**, tomando como referencia conceptual el reconocimiento de responsabilidad penal de las personas jurídicas.

El Código Penal actual sigue sirviendo

Pese a estos nuevos desafíos, Delgado rechazó la idea de que el Derecho penal vigente haya quedado inutilizado.

En los **delitos de resultado**, buena parte de los tipos penales existentes siguen siendo aplicables.

«**Ahora mismo se podría actuar con los tipos penales actuales**», sostuvo.

La razón es que muchos delitos castigan una conducta o la producción de un resultado, pero no exigen que se haya empleado un medio determinado para provocarlo.

Si una persona utiliza IA para estafar, atacar la libertad sexual o causar un daño, el empleo de esta tecnología no impide aplicar necesariamente el delito ya existente.

Otra cuestión distinta es la dificultad para **investigar, identificar al autor y obtener pruebas**.

Delgado sí apuntó, no obstante, hacia posibles reformas relacionadas con la **suplantación de identidad y los *deepfakes***.

«No tenemos que esperar a que nos cause un daño»

La tesis más contundente de su intervención llegó al final.

El magistrado sostuvo que el Derecho penal no puede limitarse siempre a reaccionar una vez producido el perjuicio.

«**No tenemos que esperar a que nos cause un daño**», afirmó.

Su propuesta pasa por recurrir en determinados supuestos a los llamados **delitos de peligro**, que permiten intervenir penalmente cuando existe un riesgo suficientemente grave, aunque todavía no se haya materializado un daño concreto.

«**Si nosotros sabemos que, por probabilidad, existe un riesgo importante de que una determinada conducta, de que un determinado sistema de IA me va a causar un daño, no tengo que esperar a que se me cause daño. Tengo que prohibirlo**», sostuvo.

Puso como ejemplo los sistemas de IA creados específicamente para generar **falsos desnudos de personas reales**.

«**Esa prohibición de tener un sistema de IA generativa que solamente se dedique a desnudar a personas tiene que ser reforzada con algo**», afirmó.

Y remató con una expresión deliberadamente directa: «**Y el mejor refuerzo que se me ocurre es a la puñetera cárcel como lo hagas**».

El papel del «Compliance»

Distinta es la situación de los **sistemas de alto riesgo** que pueden generar beneficios pero también peligros relevantes.

En esos casos, explicó Delgado, deben establecerse requisitos estrictos de funcionamiento y mecanismos que garanticen su cumplimiento.

«**Compliance. Cumplimiento, estoy hablando de eso**», resumió.

El magistrado defendió la necesidad de «**fortalecer todos los sistemas de Compliance ligados al sistema de inteligencia artificial**», especialmente en aquellas compañías que desarrollen o utilicen sistemas de alto riesgo.

Y dejó abierta la última gran incógnita.

«**Tenemos que reflexionar sobre cuál es la respuesta penal que tenemos que dar**» cuando una inteligencia artificial funciona de manera autónoma y el resultado producido ya no puede atribuirse claramente a una persona humana.

Ese, más que un problema de ciencia ficción, será uno de los grandes debates que tendrá que afrontar el Derecho penal a medida que la inteligencia artificial gane autonomía y capacidad de decisión.

Fuente: El magistrado Delgado advierte: la inteligencia artificial multiplica el delito y desdibuja quién debe responder penalmente - Confilegal