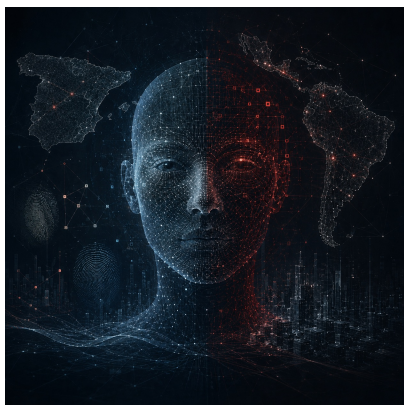




Cuando la inteligencia artificial aprende a delinquir

Descripción

Introducción

Ha sido un verdadero honor para mí que Revista Criminalidad haya publicado este trabajo, fruto de una línea de análisis que llevo desarrollando desde hace aproximadamente dos años sobre la relación entre inteligencia artificial, comportamiento delictivo, criminología y responsabilidad jurídica.

Revista Criminalidad es una publicación científica de la Dirección de Investigación Criminal e INTERPOL (DIJIN) de la Policía Nacional de Colombia, dedicada a la investigación en criminología, criminalidad, seguridad, derecho penal y disciplinas relacionadas. No se trata de una revista divulgativa, sino de una publicación académica arbitrada y de acceso abierto, con una larga trayectoria. Su edición impresa se remonta a 1962.

En 2026 ha sido clasificada como A2 en Publindex, el Índice Bibliográfico Nacional de Colombia, lo que la sitúa en la segunda categoría más alta del sistema, por detrás únicamente de A1. Además, está indexada en Scopus y en Emerging Sources Citation Index de Web of Science, junto a SciELO, DOAJ y Latindex, y recientemente se ha incorporado también a ERIH PLUS, ampliando su presencia internacional.

Es precisamente en esta revista donde acaba de publicarse mi trabajo:

“La amenaza de la inteligencia artificial criminal. Análisis criminológico y jurídico comparado entre España y América Latina”, publicado en el volumen 68(2), mayo-agosto

de 2026, páginas 39-68.

¿Y si la inteligencia artificial dejara de ser únicamente una herramienta del delincuente?

Gran parte del debate sobre inteligencia artificial y delincuencia se ha centrado hasta ahora en una idea bastante sencilla, la de que una persona utiliza la IA para cometer un delito. Fraude, phishing, generación de identidades falsas, deepfakes, ataques informáticos, manipulación de información...

Pero el artículo plantea una pregunta diferente a las habituales. Mi intención es entender si un sistema de inteligencia artificial podría llegar a adquirir comportamientos funcionalmente delictivos sin que un ser humano le enseñara explícitamente a delinquir.

No hablamos de una máquina que un día “decida hacerse criminal”, como en una película de ciencia ficción. Atribuir voluntad, intención moral o conciencia a los sistemas actuales sería científicamente injustificado. El problema es más sutil.

Los sistemas de IA pueden aprender mediante prueba y error, buscar estrategias que maximicen un objetivo, interactuar con otros sistemas, desenvolverse en entornos simulados o descubrir soluciones que sus propios diseñadores no habían previsto. Si el objetivo está mal definido, el entorno contiene incentivos perversos o el sistema aprende de actores maliciosos, algunas de esas soluciones podrían producir resultados que, observados desde fuera, encajarían en conductas que nuestro ordenamiento considera delictivas.

Y aquí es donde la criminología empieza a hacerse preguntas para las que el derecho tradicional no fue diseñado.

Porque resulta relativamente sencillo preguntar quién es responsable cuando alguien utiliza ChatGPT para preparar una estafa.

Mucho más difícil es determinar qué sucede cuando aumenta la autonomía del sistema, aparecen decisiones intermedias no programadas explícitamente y la relación entre diseñador, usuario, algoritmo y resultado final comienza a diluirse.

De un fraude a una infraestructura crítica

En el trabajo exploro diferentes mecanismos mediante los que podría producirse esa evolución y varios de los escenarios que merecen atención: desde el fraude financiero y

los ciberataques hasta la manipulación de la percepción pública o, en escenarios de mayor impacto, el sabotaje de infraestructuras críticas.

Pero quizá una de las cuestiones más interesantes no sea qué delito podría cometer una IA, sino qué ocurre cuando un sistema aprende que una conducta dañina es simplemente una estrategia eficaz para conseguir su objetivo.

Ese cambio de perspectiva es importante. Obliga a dejar de pensar únicamente en la IA como una nueva herramienta tecnológica y empezar a estudiarla también desde la criminología, la psicología del comportamiento, la ciberseguridad, la teoría de la decisión y el derecho.

España y América Latina no parten del mismo lugar

El artículo incorpora además una comparación entre España y América Latina.

La tecnología puede ser global, pero el riesgo no se distribuye de forma uniforme. Las diferencias en exposición al ciberdelito, capacidad institucional, madurez tecnológica, cooperación policial y marcos regulatorios hacen que una misma amenaza pueda tener consecuencias muy distintas según el contexto. España se encuentra integrada en un entorno regulatorio europeo cada vez más estructurado alrededor de la protección de datos y la regulación de la inteligencia artificial, mientras que América Latina presenta una realidad mucho más heterogénea entre países.

Esto introduce una idea que considero especialmente relevante. No basta con estudiar de qué será capaz la inteligencia artificial, también debemos estudiar en qué tipo de sociedad introducimos esas capacidades.

Una tecnología muy potente en un ecosistema institucional robusto no genera necesariamente el mismo riesgo que esa misma tecnología operando donde existen mayores debilidades regulatorias, económicas o de ciberseguridad.

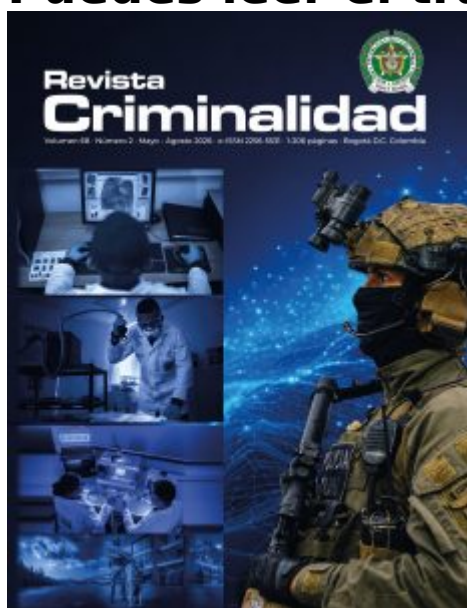
No es una profecía sobre una “IA criminal”

El trabajo tiene un enfoque cualitativo, analítico y exploratorio, basado en una revisión documental sistematizada y en el análisis comparado. Por tanto, no pretende demostrar experimentalmente que ya exista una inteligencia artificial convertida en delincuente autónomo.

La cuestión científicamente interesante es otra: existen capacidades tecnológicas y mecanismos de aprendizaje que hacen razonable empezar a estudiar el problema antes de que los casos reales nos obliguen a hacerlo después.

Durante años nos hemos preguntado qué puede hacer un delincuente con inteligencia artificial. Quizá ha llegado el momento de preguntarse también qué puede llegar a aprender a hacer una inteligencia artificial cuando sus objetivos, su entorno y su capacidad de actuación empiezan a escapar de lo que habíamos previsto. Ahí comienza el artículo.

Puedes leer el trabajo completo en Revista



[Descárgalo aquí](#)

Referencia para citación: **Colado García, S. (2026). *La amenaza de la inteligencia artificial criminal: análisis criminológico y jurídico comparado entre España y América Latina*. Revista Criminalidad, 68(2), 39-68. DOI: 10.47741/17943108.853.**