

Opinión | La abogacía ante el sesgo algorítmico en la investigación penal: por qué la abogacía debe estar presente en el diseño de la IA

Jorge Navarro Massip, vicedecano del ICAB, advierte de que la inteligencia artificial aplicada a la investigación policial debe someterse a garantías reforzadas de transparencia, supervisión humana y respeto al derecho de defensa.

«No podemos conocer la realidad tal como es en sí, sino solo tal como se nos aparece» – Immanuel Kant.

El derecho de defensa exige que la persona afectada pueda conocer, cuestionar y contradecir los elementos utilizados en su contra.

Con el uso de la IA tenemos que adelantarnos, ir al paso previo. No basta con informar que un sistema ha generado una información que se contiene en un atestado o informe policial.

La defensa debe poder saber cuáles son las instrucciones dadas y la finalidad del sistema, como se alimentaron los datos de entrenamiento de la herramienta, qué factores influyeron en la predicción, qué control, supervisión y verificación humana se aplicó, entre otras múltiples garantías.

Lo anterior no es más que aplicar una novedosa vertiente del principio de contradicción -como manifestación del derecho de defensa- en el sistema de la herramienta que va a generar esos informes o atestados.

Los cuerpos policiales de toda Europa están incorporando sistemas de inteligencia artificial. La situación es sugerente: más eficiencia y más eficacia.

Pero detrás de esa poderosa narrativa tecnológica hay una cuestión jurídica básica que apenas se debate: **¿quién y cómo se garantiza que esos sistemas no vulneren el derecho de defensa, el derecho a la presunción de inocencia o el derecho a un juicio justo?**

La Inteligencia Artificial puede y debe asistir en la investigación penal, pero no puede olvidar ni prescindir de las garantías que legitiman la intervención del Estado e inspiran el proceso penal.

El informe *Sesgos de la IA en las fuerzas del orden: una guía práctica*, publicado recientemente por el Observatorio del Laboratorio de Innovación de Europol reconoce que los sesgos de la inteligencia artificial pueden introducirse en cualquier fase del ciclo de vida del

sistema: desde el momento en que se decide qué problema se quiere resolver, pasando por la selección y el etiquetado de los datos, hasta la implementación y el uso operativo.

Y advierte -y reconoce humildemente- que, en el ámbito policial, determinados sesgos pueden traducirse en vigilancia excesiva, identificaciones injustificadas, detenciones e incluso en enjuiciamientos inadecuados.

La pregunta sobre las garantías no es retórica. Hoy en día la respuesta no solo es jurídica, también es tecnológica. El algoritmo incide directamente sobre derechos fundamentales de las personas. El propio Reglamento define como de alto riesgo los sistemas de IA utilizados por las autoridades judiciales para la investigación y la interpretación jurídica.

La inteligencia artificial no es neutral

Es importante cuestionar la idea de neutralidad u objetividad de la IA. Los sistemas de inteligencia artificial, especialmente los basados en aprendizaje automático, descubren patrones ocultos en los datos.

Pero no instrumentos “asépticos“, contienen sesgos. Son sistemas sociotécnicos: interactúan componentes técnicos (algoritmos, datos...) y componentes sociales (diseño por personas, normas, contextos de uso, etc.)

El propio informe de Europol lo explica con claridad. **En la fase de diseño**, la forma en que se formula el problema **ya puede incorporar sesgos**.

En la **fase de adquisición de datos** aparecen otros sesgos. Los datos policiales pueden reflejar desigualdades históricas, prácticas de vigilancia selectiva, discriminación institucional o ideológica.

A esto se suman los sesgos de etiquetado (cuando las personas que anotan los datos incorporan sus propios prejuicios), los sesgos de medición (cuando una variable actúa como sustituto indirecto de una característica protegida, como el código postal que puede condicionar el nivel socioeconómico o el origen étnico) o los sesgos de implementación (cuando el modelo se utiliza en un contexto distinto de aquel para el que fue diseñado).

«Ningún sistema de inteligencia artificial puede convertir una predicción estadística en una presunción de culpabilidad, ni puede privar a la defensa de los medios necesarios para conocer, cuestionar y contradecir los elementos utilizados contra una persona. Una vez más, sin Abogacía no hay Estado de Derecho».

Un sistema entrenado con determinados datos puede concluir (¿aprender?) que determinados barrios, nacionalidades o perfiles sociales están asociados con la delincuencia.

La inteligencia artificial puede concluir con un silogismo que aparentemente sea lógico, puede contener un sesgo de confirmación, puede determinar y reproducir una pauta introducida preexistente, o simplemente optar por la hipótesis más sencilla.

En este sentido, **los juristas sabemos que la navaja de Ockham no es aceptable en un proceso penal** porque la verdad jurídica no se decide por la simplicidad de la hipótesis, sino por la existencia de una prueba suficiente y válidamente obtenida.



El Reglamento Europeo prohíbe, salvo excepciones, el uso de sistemas de inteligencia artificial para la vigilancia policial predictiva cuando se basen exclusivamente en la elaboración de perfiles de una persona física o en la evaluación de sus rasgos y características de personalidad.

Derechos fundamentales en riesgo

Cuando estos sesgos se proyectan sobre el proceso penal, se afectan las garantías estructurales del proceso. El artículo 48 de la Carta de los Derechos Fundamentales de la Unión Europea proclama que toda persona acusada se presume inocente hasta que su culpabilidad haya sido legalmente declarada.

La presunción de inocencia también exige que la prueba a valorar haya sido válidamente obtenida, no siendo aptas aquellas que se haya obtenido vulnerando derechos fundamentales.

Hoy en día, especialmente, aplica sobre la autenticidad de la prueba. Y la pregunta es: ¿este principio elemental se va a introducir en los sistemas de IA policial como una instrucción del sistema (system prompt)?

En análogo sentido ¿la herramienta valora el silencio del acusado? No hay que ser jurista para entender que no incluir estos derechos o garantías colisiona con los derechos fundamentales de la ciudadanía.

El Ministerio Fiscal, los Jueces, la defensa: todos debemos saber de forma clara y precisa la procedencia de los datos, el contexto de entrenamiento o la trazabilidad de la decisión.

La falta de información o, simplemente, la incertidumbre por parte de los juristas que participan en el proceso debe operar como un línea roja en la utilización incriminatoria del resultado algorítmico policial.

Para que una defensa sea efectiva, la defensa debe tener acceso al código fuente, a los datos de entrenamiento y a los manuales del algoritmo.

Si la defensa no puede auditar o contrastar como llegó la IA a una conclusión incriminatoria, se produce una desventaja procesal que vulnera el principio de igualdad de armas y el juicio equitativo.

El Reglamento europeo como presupuesto y garantía jurídica

Recientemente ha entrado en vigor -parcialmente- un marco jurídico vinculante: el **Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, sobre normas armonizadas en materia de inteligencia artificial**.

Este Reglamento clasifica como sistemas de alto riesgo, entre otros, aquellos destinados a ser utilizados por las autoridades encargadas de la aplicación de la ley para evaluar el riesgo de que una persona cometa un delito, detectar o identificar personas en el contexto de la investigación penal, evaluar la fiabilidad de las pruebas, elaborar perfiles de personas físicas en el curso de la investigación penal o realizar policía predictiva aplicada a personas físicas.

Esta calificación no es meramente descriptiva. Implica la aplicación de un conjunto de presupuestos jurídicos vinculantes que deben cumplirse antes del despliegue del sistema: gestión continua de riesgos, gobernanza de datos, documentación técnica, transparencia, supervisión humana, etc.

Además, exige la realización de una evaluación de impacto relativa a los derechos fundamentales antes de la puesta en servicio.

Policía predictiva: cuando la vigilancia se convierte en prueba

La policía predictiva constituye el ámbito donde de forma más intensa confluyen los riesgos descritos.

El Reglamento Europeo prohíbe, salvo excepciones, el uso de sistemas de inteligencia artificial para la vigilancia policial predictiva cuando se basen exclusivamente en la elaboración de perfiles de una persona física o en la evaluación de sus rasgos y características de personalidad.

La Guía de Europol menciona casos concretos que ilustran estos riesgos. El proyecto neerlandés “Sensing”, desarrollado en Roermond, utilizaba sensores para identificar características de vehículos y rutas asociadas a personas consideradas sospechosas de delitos de hurto.

El sistema acabó atribuyendo un riesgo mayor a ciudadanos de Europa del Este o de etnia romaní y se excluía a los neerlandeses.

También “Precobs”, utilizado en Alemania y Suiza para predecir determinados delitos a partir de hechos pasados. Las fuerzas policiales dejaron de utilizarlo debido a la insuficiencia de datos para realizar predicciones fiables.

Por qué la abogacía debe participar activamente en el diseño

La Abogacía no puede limitarse a impugnar o cuestionar los resultados cuando ya se ha producido la investigación o la detención, como se hace en la actualidad con los ordinarios medios de prueba y la clásica fase de investigación policial y judicial.

El escenario es otro, y los actores también. La IA es la estrella invitada.

La Abogacía, como institución ajena al Estado, cuya intervención y participación exige el propio Estado para garantizar el derecho de defensa mediante la asistencia letrada, debe participar desde el inicio.

La Abogacía debe anticipar la afectación de derechos, delimitar las finalidades legítimas, atender a los principios de proporcionalidad y necesidad y un sinnúmero de aspectos.

Por ello, hay que participar en la preparación de protocolos policiales, en la contratación pública de sistemas de IA, en la definición de las finalidades, en la selección de proveedores, en la elaboración de pliegos técnicos y en la creación de procedimientos de supervisión.

La transparencia y el conocimiento es más necesario que nunca. Las auditorías de estos sistemas deben incluir una dimensión jurídica que se pregunte si el sistema tiene una base garantista, si la finalidad es concreta y legítima, si se tratan más datos de los necesarios, qué sucede con esos datos, si el resultado se utiliza para generar sospecha, si la defensa puede acceder a la información relevante, si se preserva la presunción de inocencia, etc.

Es imprescindible que la transparencia tenga una dimensión en el proceso penal y en la investigación policial. **Cuando un resultado algorítmico haya participado de una actuación policial o judicial, la defensa debe poder conocer cómo se gestó esa información para, en su caso, poder cuestionarla.**

La conclusión es sencilla: ningún sistema de inteligencia artificial puede convertir una predicción estadística en una presunción de culpabilidad, ni puede privar a la defensa de los medios necesarios para conocer, cuestionar y contradecir los elementos utilizados contra una persona. Una vez más, sin Abogacía no hay Estado de Derecho.

<https://confilegal.com/20261001-ia-policial-derecho-defensa-jorge-navarro-massip/>