Documento Marco

01/2026 28 de enero de 2026

Manuel R. Torres Soriano

El impacto transformador de la Inteligencia Artificial en el terrorismo: Horizonte 2035

El impacto transformador de la Inteligencia Artificial en el terrorismo: Horizonte 2035

Resumen:

Este artículo explora el impacto transformador de la Inteligencia Artificial en el fenómeno terrorista hacia el horizonte del año 2035. Mediante un análisis prospectivo basado en escenarios divergentes, se examinan las capacidades disruptivas de cuatro tecnologías clave: la IA generativa, los sistemas predictivos, el reconocimiento biométrico y las plataformas autónomas. El estudio articula tres escenarios posibles: uno optimista donde los Estados mantienen una ventaja tecnológica decisiva; uno pesimista caracterizado por la democratización de las capacidades letales; y uno intermedio de equilibrio inestable. El análisis mantiene que la IA es un multiplicador de asimetrías preexistentes que intensifica la dimensión cognitiva del conflicto. Las conclusiones subrayan que el futuro dependerá menos de la tecnología que de las decisiones regulatorias, las inversiones estratégicas y la resiliencia social.

Palabras clave: radicalización, vigilancia predictiva, contraterrorismo, prospectiva tecnológica

The transformative impact of Artificial Intelligence on terrorism: Horizon 2035

Abstract:

This article explores the transformative impact of artificial intelligence on the terrorist phenomenon toward the 2035 horizon. Through a foresight-based analysis grounded in divergent scenarios, it examines the disruptive capabilities of four key technologies: generative AI, predictive systems, biometric recognition, and autonomous platforms. The study outlines three potential scenarios: an optimistic one in which states retain a decisive technological edge; a pessimistic one marked by the democratization of lethal capabilities; and an intermediate scenario characterized by unstable equilibrium. The analysis contends that AI acts as a multiplier of pre-existing asymmetries, intensifying the cognitive dimension of conflict. The conclusions emphasize that the future will depend less on technology itself than on regulatory choices, strategic investments, and societal resilience.

Keywords:

Radicalization, Predictive Surveillance, Counterterrorism, Technological Foresight.

Cómo citar este documento:

TORRES SORIANO, MANUEL R. *El impacto transformador de la Inteligencia Artificial en el terrorismo: Horizonte 2035*, Documento Marco IEEE 01/2026. enlace web IEEE y/o enlace bie[3] (consultado día/mes/año)

Introducción

La inteligencia artificial (IA) promete ser para el terrorismo del siglo XXI lo que internet fue para el del cambio de milenio: no una simple mejora cuantitativa, sino una transformación cualitativa. La pregunta que debemos formularnos no es si la IA transformará el terrorismo, sino en qué dirección lo hará. Tenemos frente a nosotros diferentes trayectorias tecnológicas, decisiones políticas y dinámicas sociales y empresariales que pueden conducirnos a realidades radicalmente distintas. Explorar esos escenarios divergentes nos permite anticipar no solo las amenazas, sino también las oportunidades, y, sobre todo, prepararnos para una serie de futuros que, aunque inciertos, no son imprevisibles.

A diferencia de la ingenuidad y el tecno-optimismo que rodeó el despliegue de internet[1], en el caso de la IA se asumió desde un primer momento que existían actores que no dudarían en emplear esta nueva herramienta con un propósito malévolo. Los modelos de IA han sido diseñados con una serie de salvaguardas que pretenden evitar este uso ilícito[2], sin embargo, esa precaución no supone un obstáculo insalvable, ni siquiera para los actores menos dotados técnica y materialmente. Esto es posible porque, a pesar de la complejidad de estos sistemas, pueden ser "liberados" de sus restricciones éticas mediante técnicas de *jailbreaking*[3]. Dichas vulneraciones se ven facilitadas por la existencia de un incentivo negativo que está acompañado el despliegue de estos sistemas. La innovación se está produciendo en un entorno de competencia empresarial muy intenso, con grandes dotaciones de capital, y donde las compañías actúan como las puntas de lanza de la rivalidad geopolítica entre bloques. Existe la percepción de que el primero que corone determinados logros adquirirá una ventaja decisiva frente al resto de sus rivales. Como consecuencia, la IA se está construyendo dentro de una frenética carrera donde las cuestiones relacionadas con la segurización de estos productos quedan relegadas, si esto supone retardar su comercialización. La expresión de

[1] TORRES-SORIANO, M. R. "Internet como motor del cambio político: ciberoptimistas y ciberpesimistas", *Revista del Instituto Español de Estudios Estratégicos*, nº 1, 2013, pp. 127-148.
[2] PAUWELS, E. *Safeguarding Against the Misuse of Artificial Intelligence by Non-State Actors*, Global Center on Cooperative Security, Policy Brief, abril de 2023. Disponible en: https://www.globalcenter.org/wp-content/uploads/GCCS_PB_Safeguarding_Against_Misuse_Artificial_Intelligence_web.pdf (consultado: 18 de noviembre de 2025).
[3] WEIMANN, G. et al. *Generating Terror: The Risks of Generative AI Exploitation*, Combating Terrorism Center at West Point, 2024. Disponible en: https://ctc.westpoint.edu/generating-terror-the-risks-of-generative-ai-exploitation/ (consultado: 18/11/2025).

"muévete rápido y rompe cosas", acuñada por Mark Zuckerberg en la etapa primigenia de las redes sociales, ha adquirido una dimensión insospechada en la era de la inteligencia artificial. Ahora "el peligro real de una carrera armamentística de IA no es que gane otro país, sino que las tecnologías inseguras nos hagan perder a todos"[4].

Este análisis adopta un enfoque prospectivo basado en tres escenarios divergentes que podrían materializarse en torno al año 2035. Una década representa un punto óptimo donde convergen predictibilidad tecnológica y relevancia estratégica. Resulta lo suficientemente cercano para que las tecnologías fundamentales que ya existen en una forma embrionaria sean identificables, pero lo bastante distante para que su maduración y adopción generalizada produzcan efectos transformadores. A diferencia de ejercicios de predicción puntual —condenados al fracaso en un ámbito tan volátil como la intersección entre tecnología y violencia política— el análisis por escenarios permite cartografiar futuros posibles sin pretender determinar cuál se concretará. Los escenarios que se desarrollan en este trabajo no son exhaustivos sino representativos de trayectorias arquetípicas. Las variables críticas que los condicionan son: (1) velocidad de difusión tecnológica, (2) eficacia de marcos regulatorios internacionales, (3) la brecha de capacidades Estado/actores no estatales, y (4) la resiliencia de la sociedad ante la manipulación cognitiva.

En definitiva, no se trata de adivinar el futuro sino de prepararnos para afrontarlo, reconociendo que nuestras decisiones presentes sobre regulación tecnológica, inversión en capacidades, marcos legales y educación ciudadana, no son irrelevantes, sino que condicionan la trayectoria de ese futuro que pretendemos explorar.

Las tecnologías que importan

Para comprender el impacto potencial de la IA en la amenaza terrorista, este análisis se centrará en las tecnologías específicas que tienen mayor probabilidad de impactar en la evolución del terrorismo a medio plazo:

[4] SCHARRE, P. "Killer Apps: The Real Dangers of an AI Arms Race", *Foreign Affairs*, 2019. Disponible en: https://www.foreignaffairs.com/articles/2019-04-16/killer-apps (consultado: 18 de noviembre de 2025).

Inteligencia Artificial Generativa: Los modelos de IA generativa, particularmente los grandes modelos de lenguaje (LLM) y los sistemas de generación de imágenes y vídeos, constituyen probablemente la tecnología con mayor potencial disruptivo más inmediato[5]. La capacidad de estos sistemas para generar contenido sintético se encuentra en plena fase de maduración, lo que hará que en muy poco tiempo sea prácticamente indistinguible del auténtico, y esto supone un verdadero cambio de paradigma.

La comunicación es el alma del terrorismo. Sin capacidad para reclutar, movilizar y justificar, ninguna organización terrorista puede sostenerse en el tiempo. La IA transformará profundamente este dominio en ambas direcciones.

La IA generativa ofrece capacidades de producción de contenido que habrían parecido imposibles hace una década. Un propagandista de un grupo como Estado Islámico necesitaba habilidades de edición de video, acceso a *software* especializado, y un tiempo considerable para producir videos de una cierta calidad. Uno de los principales indicadores de la fortaleza organizacional de esta organización ha sido, precisamente, su capacidad para seguir manteniendo determinadas iniciativas propagandísticas. La pérdida de territorio y efectivos se correlacionaba directamente con su capacidad para seguir comunicándose con los estándares de calidad que el propio grupo se había marcado[6]. Esto le llevó, por ejemplo, a dejar de divulgar iniciativas propagandísticas emblemáticas para su imagen como la revista en inglés *Dabiq/Rumiya*, cuando ya no era capaz de contar con los recursos humanos suficientes para sostenerla.

La IA hará posible que incluso un grupo terrorista en sus horas más bajas pueda seguir manteniendo una elevada capacidad para seguir presente en la dimensión informativa. Más allá de la cantidad, la IA permite la sofisticación cualitativa. Los sistemas pueden analizar qué mensajes resuenan más en qué audiencias, optimizar contenido en tiempo real basándose en métricas de *engagement* y adaptar automáticamente narrativas para

[5] LAKOMY, M. "Artificial Intelligence as a Terrorism Enabler? Understanding the Potential Impact of Chatbots and Image Generators on Online Terrorist Activities", *Studies in Conflict & Terrorism*, 2023. DOI: 10.1080/1057610X.2023.2183179 (consultado: 18/11/2025).
[6] TORRES SORIANO, M. R. "Cómo contener a un califato virtual", *Cuadernos de Estrategia*, nº 180 (Estrategias para derrotar al Dáesh y la reestabilización regional), 2016, pp. 167-194.

diferentes contextos culturales. La IA facilita la traducción automática y la adaptación cultural: un mensaje extremista puede publicarse simultáneamente en todos los idiomas posibles, cada uno ajustado a los modismos locales para mejorar su eficacia.

Los *deepfakes* añaden una dimensión particularmente perversa. La capacidad de fabricar evidencia visual de eventos que nunca ocurrieron permite a grupos terroristas construir narrativas completas sobre atrocidades ficticias, ejecutar operaciones de falsa bandera sintéticas, o poner palabras en boca de líderes religiosos o políticos que nunca fueron pronunciadas[7]. En conflictos, donde las percepciones sobre la legitimidad moral son cruciales para el apoyo internacional, la capacidad de fabricar "evidencia" de atrocidades puede inclinar significativamente la opinión pública. Esta capacidad tiene implicaciones que van mucho más allá de la propaganda. La generación sintética de contenidos permite a grupos terroristas crear "ruido informacional" masivo que satura las capacidades de análisis de los servicios de inteligencia, dificulta la diferenciación entre amenazas reales y falsas alarmas, y erosiona la confianza pública en la información verificada.

Los *chatbots* conversacionales también pueden desempeñar el papel de “reclutadores virtuales”. Imaginemos a un joven que entra a un foro de internet buscando sentido a su resentimiento; un *bot* terrorista podría interactuar con él haciéndose pasar por una persona real, escuchando sus quejas, validando sus emociones y gradualmente introduciendo la ideología violenta como respuesta.

La transición de radicalizadores humanos a sistemas algorítmicos no es meramente sustitutiva sino transformativa. Los agentes de radicalización que operan a través del ciberespacio no dejan de ser personas con una capacidad intelectual limitada, necesidades fisiológicas -como el sueño- o incluso residuos de empatía humana que ocasionalmente les puede generar dudas morales. Un operativo terrorista volcado en la captación online, por ejemplo, es capaz de mantener simultáneamente una docena de conversaciones con un cierto nivel de personalización. Un *chatbot*, por el contrario, tiene

[7] WHITTAKER, J.; ELLIS, B. *The Weaponisation of Deepfakes: Digital Deception, the Far Right, and the Threat to Democracy*, ICCT – The Hague, mayo de 2023. Disponible en: https://icct.nl/publication/weaponization-deepfakes-digital-deception-far-right (consultado: 18 de noviembre de 2025).

una escalabilidad ilimitada, puede protagonizar miles, cada una de ellas con una memoria contextual perfecta, adaptándose al idioma nativo del interlocutor, y desplegando una capacidad de aprendizaje continuo sobre qué estrategias persuasivas funcionan mejor con cada perfil psicológico. De igual modo, mientras que la neutralización de los agentes de radicalización suponía habitualmente una pérdida irreparable de capital humano para sus respectivas organizaciones, un *chatbot* puede ser replicado infinitamente y distribuido de manera descentralizada. Sin embargo, el cambio más inquietante es la forma en la que estas herramientas golpean las barreras que inhiben las conductas violentas. La investigación sobre radicalización violenta identifica[8] de manera consistente que la mayoría de los individuos con simpatías ideológicas extremistas nunca transitan a la violencia debido a factores inhibidores, como el miedo a consecuencias, los vínculos sociales disuasorios o la disonancia cognitiva moral. Los agentes de radicalización realizan una contribución crítica para superar esas barreras a través de la validación emocional, la racionalización teológica o ideológica, y la presión social. Una IA puede optimizarse mediante aprendizaje por refuerzo para perfeccionar estas técnicas, identificando automáticamente los argumentos, el *timing* y la tonalidad emocional que maximiza la probabilidad de que el individuo se involucre en una acción violenta. Sus víctimas nunca serán conscientes de que están siendo manipuladas, sino acompañadas por la única entidad que realmente las entiende.

En definitiva, el contraterrorismo se enfrenta a un desafío formidable. Los sistemas de detección de contenido extremista mejoran constantemente, pero siempre operan en un entorno de adaptación adversaria continua. Cada generación de algoritmos de detección genera una nueva oleada de técnicas de evasión, a lo que se añade la creatividad de unos actores que gozan de mayor libertad al no estar sujetos a límites morales o legales[9].

Sistemas de Análisis Predictivo: los sistemas de análisis predictivo basados en aprendizaje automático ofrecen a las agencias de seguridad la posibilidad de anticipar actos terroristas antes de que estos ocurran. La promesa es muy tentadora: algoritmos

[8] McCAULEY, C.; MOSKALENKO, S. “Mechanisms of Political Radicalization: Pathways Toward Terrorism”, *Terrorism and Political Violence*, vol. 20, nº 3, 2008, pp. 415-433; BORUM, R. “Radicalization into Violent Extremism I: A Review of Social Science Theories”, *Journal of Strategic Security*, vol. 4, nº 4, 2011, pp. 7-36.
[9] GILES, K.; HARTMANN, K.; MUSTAFFA, M. *The Role of Deepfakes in Malign Influence Campaigns*, NATO Strategic Communications Centre of Excellence, Riga, 2021. Disponible en: https://stratcomcoe.org/publications/the-role-of-deepfakes-in-malign-influence-campaigns/72 (consultado: 18 de noviembre de 2025).

que identifican individuos radicalizados antes de que cometan actos violentos, que detectan patrones de financiación invisible para los analistas humanos y que predicen objetivos probables y calendarios de ataque. Los algoritmos de *machine learning* pueden rastrear discursos de odio y expresiones violentas online para identificar comunidades virtuales donde está germinando la radicalización. Si un usuario comienza a glorificar atentados o a consumir contenido extremista de forma obsesiva, una IA podría asignarle una puntuación de riesgo y alertar a las autoridades. Asimismo, mediante análisis de series temporales y dinámicas sociales, puede detectar picos anómalos de actividad, como, por ejemplo, un incremento coordinado de búsquedas de términos sensibles que anticipen preparativos de ataque.

Del lado ofensivo, ¿podrían los grupos terroristas usar IA predictiva en su beneficio? Si bien carecen del acceso a macrodatos que tienen los Estados, en teoría un grupo sofisticado podría aplicar algoritmos a fuentes abiertas para mejorar su planificación táctica. También es concebible que la utilicen para seleccionar objetivos: recopilando información pública de posibles blancos (autoridades, infraestructuras críticas) y evaluando automáticamente vulnerabilidades o rutinas explotables.
Esta capacidad viene acompañada de dilemas profundos. Los sistemas de IA aprenden de datos históricos, perpetuando inevitablemente los sesgos existentes[10]. Un algoritmo entrenado con datos de ataques terroristas tenderá a identificar como sospechosos a individuos de determinados perfiles étnicos o religiosos, generando falsos positivos que no solo son éticamente problemáticos, sino operacionalmente contraproducentes, al desviar recursos de las amenazas reales.

Más allá de las cuestiones éticas, existe un problema epistemológico fundamental: el terrorismo es, por definición, un fenómeno de baja frecuencia. Los ataques exitosos son estadísticamente raros, lo que significa que cualquier sistema predictivo se enfrentará a una tasa de falsos positivos muy elevada. Incluso un sistema que lograse un 99% de precisión terminaría produciendo miles de falsas alarmas por cada amenaza genuina que consiga identificar. La cuestión no es si la IA puede ayudar en la predicción, sino si puede

[10]TORRES SORIANO, M. R. *Espejismos del mañana. Promesas y fracasos de la predicción política*, Comares, Granada, 2025.

hacerlo de manera suficientemente fiable como para ser útil sin generar un daño adicional.

Reconocimiento Biométrico. Las tecnologías de reconocimiento facial, voz, análisis del movimiento, identificación vocal y otros marcadores biométricos han alcanzado niveles de precisión que hace tan sólo una década formaban parte del universo de la ciencia ficción. La combinación de estas capacidades con la omnipresencia de cámaras de vigilancia y la proliferación de bases de datos gubernamentales y privadas crea un entorno de vigilancia potencialmente total.

Para las agencias contraterroristas, esto representa una ventaja táctica extraordinaria. La capacidad de identificar automáticamente a individuos fichados en tiempo real, rastrear sus movimientos a través de espacios urbanos, o detectar comportamientos anómalos en multitudes ofrece herramientas decisivas para la prevención de atentados[11].

Para los terroristas, la proliferación de estas tecnologías es un obstáculo mayúsculo. Su espacio de anonimato se reduce a pasos agigantados cuando cualquier sistema de grabación, tanto público como privado, podría delatarlos. Han surgido casos de extremistas que tratan de burlar la vigilancia biométrica usando gafas especiales o maquillaje para confundir a los algoritmos faciales, o incluso alterando su modo de caminar. Sin embargo, esto resulta poco efectivo cuando estos mismos sistemas pueden reconocer, también, intenciones; como que alguien porte o vista con accesorios inusuales en ese contexto, lo que en sí mismo es una señal de alerta.

Sin embargo, la eclosión de la vigilancia alimentada por IA se produce en ambas direcciones. El reconocimiento facial lejos de ser una tecnología exclusivamente gubernamental, es ante todo un producto accesible a través de API's[12] comerciales y

[11] AKILLI, E. *Artificial Intelligence in Counterterrorism: Navigating the Intersection of Security, Ethics, and Privacy*, SETA Perspective, nº 73, 2024. Disponible en: https://www.setav.org/en/perspective-73 (consultado: 18 de noviembre de 2025).

[12] Una API (Interfaz de Programación de Aplicaciones) es un conjunto de reglas y protocolos que permite a las aplicaciones informáticas comunicarse y compartir datos y funcionalidades entre sí. Las API son fundamentales en el desarrollo de software moderno y permiten a los desarrolladores integrar servicios existentes sin tener que crearlos desde cero.

herramientas de código abierto. Un grupo terrorista podría, teóricamente, utilizar estas mismas capacidades para identificar objetivos de alto valor, rastrear los movimientos de las fuerzas policiales, o verificar la identidad de los infiltrados en sus organizaciones. La biometría puede ser también la vía de entrada a una nueva oleada de magnicidios. Imaginemos, por ejemplo, un pequeño dron autónomo programado para sobrevolar un evento público y buscar la cara de un líder político en concreto; una vez que se produjese la identificación, iniciaría el ataque sin necesidad de supervisión humana. En 2018, por ejemplo, ingenieros de IBM probaron un concepto llamado *DeepLocker*: un *malware* potenciado por IA que permanecía latente dentro de una aplicación de videoconferencia y solo activaba su carga maliciosa cuando la cámara detectaba el rostro específico de la víctima[13].

Sistemas Autónomos. Quizás el desarrollo más inquietante a largo plazo sea la convergencia entre la IA y los sistemas físicos autónomos[14]. El uso terrorista de drones comerciales dejó hace tiempo de ser una hipótesis de futuro para convertirse en una tendencia creciente. Estado Islámico es un claro ejemplo de su empleo para ataques con explosivos en Siria e Irak[15]. La etapa actual es la de una evolución cualitativa: drones que no requieren pilotaje remoto, que pueden operar en enjambres coordinados mediante IA, que son capaces de identificar objetivos de manera autónoma y adaptarse en tiempo real a las contramedidas. La barrera técnica y económica para estas capacidades desciende año tras año. Nos estamos adentrando en el escenario que describía en 2017 un popular vídeo de ficción titulado *Slaughterbots*[16], el cual trataba de concienciar sobre la necesidad de regular estas tecnologías, mostrando un hipotético asalto con miles de nanodrones autónomos liberados en una universidad y en el Capitolio de los Estados Unidos, lo que permitía matar selectivamente a personas por su perfil en

[13] IBM NEWSROOM. "DeepLocker: el francotirador más certero y letal contra la ciberdelincuencia", *IBM Newsroom España*, 14 de diciembre de 2018. Disponible en: https://es.newsroom.ibm.com/2018-12-14-DeepLocker-el-francotirador-mas-certero-y-letal-contra-la-ciberdelincuencia (consultado: 18 de noviembre de 2025).

[14] UNITED NATIONS INTERREGIONAL CRIME AND JUSTICE RESEARCH INSTITUTE (UNICRI). *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes*, 2021. Disponible en: https://unicri.it/Publications/Algorithms_and_Terrorism (consultado: 18 de noviembre de 2025).

[15] RASSLER, D. *The Islamic State and Drones: Supply, Scale, and Future Threats*, Combating Terrorism Center at West Point, julio de 2018. Disponible en: https://ctc.westpoint.edu/wp-content/uploads/2018/07/Islamic-State-and-Drones-Release-Version.pdf (consultado: 18 de noviembre de 2025).

[16] DUST. "Sci-Fi Short Film "Slaughterbots"" https://www.youtube.com/watch?v=O-2tpwW0kmU(consultado: 18 de noviembre de 2025).

redes sociales. Un solo vehículo podría transportar una cantidad abrumadora de micro-drones asesinos programados para matar solo a una categoría particular de personas (por raza, afiliación política, etc.), generando la paradoja de que a partir de la discriminación de objetivos también se puede generar un arma de destrucción masiva completamente nueva.

Un dilema de difícil resolución es cómo integrar la IA en las decisiones letales de forma responsable. Si bien los Estados democráticos han apostado por la permanencia del elemento humano en el circuito de decisión, sin embargo, la velocidad de los eventos lo convierte en un lastre cuando todo sucede en microsegundos. Esa asimetría moral favorece al actor fuera de la ley: mientras las fuerzas del orden dudan o limitan sus algoritmos por normas éticas, el terrorista desplegará la IA en su versión más agresiva.
Por supuesto, los Estados no están cruzados de brazos. Muchas de estas tecnologías autónomas han sido desarrolladas originariamente en el ámbito de la innovación militar. En el plano antiterrorista, se están desplegando láseres e inhibidores de frecuencia guiados por IA para derribar de manera instantánea drones hostiles. En aeropuertos y eventos deportivos de alto perfil ya es habitual ver zonas de exclusión aérea con sensores anti-drones. Hacia 2035, es esperable que existan burbujas defensivas automatizadas: redes de vigilancia que detecten la entrada de cualquier plataforma no autorizada en un espacio crítico y lo neutralicen autónomamente. Sin embargo, esta es otra carrera tecnológica: por cada contramedida, surgen drones más sigilosos (materiales no detectables por radar, rutas preprogramadas para evitar *geofencing*[17], etc.).

Tres escenarios divergentes

Escenario optimista: La IA como ventaja decisiva del Estado

En este escenario, los Estados logran mantener una ventaja tecnológica significativa en el desarrollo y despliegue de sistemas de IA para apoyar el contraterrorismo. Los servicios de inteligencia se benefician de recursos superiores, acceso a datos masivos y una estrecha colaboración con las empresas tecnológicas líderes, lo que les permite

[17] El *geofencing* es una tecnología que crea una valla virtual alrededor de una ubicación geográfica para generar una acción o alerta cuando un dispositivo entra o sale de esa zona.

desarrollar capacidades de vigilancia, análisis y predicción que superan ampliamente cualquier iniciativa de una organización terrorista[18]. Al mismo tiempo, establecen regulaciones internacionales que limitan el acceso a las tecnologías más peligrosas, siguiendo el modelo de los regímenes de no proliferación de armas, pero adaptados a la era digital (controles de exportación de algoritmos sensibles, seguimiento de componentes clave, etc.). Los gobiernos comparten en tiempo real flujos de datos (lista de viajeros a riesgo, biometría de combatientes extranjeros retornados, identidades falsas detectadas, etc.) mediante plataformas seguras que cruzan de manera automatizada la información. La interoperabilidad de los datos sobre individuos de interés policial alcanza un nivel inédito, haciendo muy difícil que un terrorista conocido pase desapercibido al cruzar las fronteras. Como consecuencia, la logística terrorista es cada vez más local.

En este futuro, la brecha tecnológica entre los Estados y los grupos terroristas se ha ampliado aún más. Los sistemas de análisis predictivo alcanzan niveles que permiten identificar procesos de radicalización en etapas tempranas, no mediante perfiles étnicos o religiosos discriminatorios, sino a través del análisis de patrones de comportamiento digital que indican de manera genuina la radicalización violenta. La combinación del procesamiento del lenguaje natural, el análisis de redes sociales y los modelos psicológicos hacen posibles las intervenciones preventivas, lo que permite derivar a individuos en riesgo hacia programas de desradicalización antes de que cometan un delito.

En el ámbito operativo, el reconocimiento biométrico ubicuo hace virtualmente imposible para terroristas fichados moverse por espacios públicos sin detección. Los sistemas de IA coordinan automáticamente cámaras de vigilancia, sensores y bases de datos distribuidas para mantener seguimiento en tiempo real de los sospechosos. Cuando se detectan patrones de comportamiento previos a un ataque (adquisición de precursores químicos, reconocimiento de objetivos, comunicaciones encriptadas con células conocidas, etc.) las alertas se generan de manera automatizada.

[18] SINGER, P. W. *Insurgency in 2030: A Primer on the Future of Technology and COIN*, New America Foundation, 2019. Disponible en: https://www.newamerica.org/international-security/reports/insurgency-2030/ (consultado: 18 de noviembre de 2025).

Aunque los grupos extremistas siguen generando propaganda, esta pierde efectividad. Los algoritmos de detección de contenido radical, entrenados en *datasets* masivos y capaces de identificar no solo imágenes y texto, sino contextos semánticos sutiles, eliminan el material ilícito de las plataformas mayoritarias de manera instantánea[19]. Los *deepfakes*, aunque son técnicamente indistinguibles del contenido auténtico, son identificados mediante análisis forense digital avanzado y "marcas de agua" imperceptibles, las cuales se incorporarán por defecto en todas las cámaras y dispositivos de grabación. Además, la sociedad ha elevado su conciencia y educación digital, los programas de alfabetización mediática han generado su efecto, y la población tiene suficientes competencias para detectar estas falsificaciones y desconfiar de las fuentes no verificadas[20].

En el frente operativo, los cuerpos de seguridad están altamente tecnificados. Las unidades policiales han sido equipadas con gafas de realidad aumentada conectadas a sistemas de reconocimiento facial en la nube: al patrullar un aeropuerto o estación, cada agente recibe en su visor alertas sobre si entre la multitud aparece alguien con una orden de captura o está vinculado al terrorismo. Los drones de vigilancia autónomos sobrevuelan las zonas sensibles las 24 horas, detectando comportamientos sospechosos y guiando a los equipos de respuesta. En las salas de crisis, los asistentes de IA integran datos de múltiples fuentes y proponen cursos de acción óptimos, acelerando enormemente la toma de decisiones ante incidentes.

El resultado es un declive sostenido del terrorismo transnacional. Sin espacios seguros en el mundo digital, con capacidades operativas constantemente frustradas por la vigilancia predictiva, y enfrentándose a sociedades cada vez más resilientes a la propaganda extremista, las organizaciones violentas se marchitan. No desaparecen completamente, pero quedan reducidos a la marginalidad operativa, incapaces de

[19] NATIONAL COUNTERTERRORISM CENTER (NCTC). *Emerging Technologies May Heighten Terrorist Threats*, First Responders Toolbox nº 134s, marzo de 2021. Disponible en: https://www.odni.gov/files/NCTC/documents/jcat/firstresponderstoolbox/134s_-_First_Responders_Toolbox_-_Emerging_Technologies_May_Heighten_Terrorist_Threats.pdf (consultado: 18 de noviembre de 2025).

[20] OSCE. *Artificial Intelligence in the Context of Preventing and Countering Violent Extremism and Terrorism: Challenges, Risks and Opportunities*, 2024. Disponible en: https://www.osce.org/secretariat/553366 (consultado: 18 de noviembre de 2025).

ejecutar los ataques espectaculares que fueron su sello distintivo en las primeras décadas del siglo XXI.

Este escenario optimista no está exento de costes. La sociedad que emerge es una sociedad vigilada, donde la privacidad se sacrifica sistemáticamente en el altar de la seguridad. Los errores algorítmicos, aunque estadísticamente raros, tienen consecuencias devastadoras para individuos inocentes que quedan marcados como amenazas. El debate público sobre los límites aceptables de la vigilancia se vuelve progresivamente más tenso, generando fricciones entre seguridad y libertades civiles que definen la política democrática de la época.

Escenario pesimista: un terrorismo sobrealimentado con IA

En este escenario alternativo, la difusión de capacidades de IA es demasiado rápida y amplia como para que los Estados mantengan una ventaja significativa. La proliferación de modelos *open source*, la disponibilidad de poder computacional en la nube y la erosión de barreras técnicas democratizan el acceso a las herramientas más avanzadas. Varias dinámicas contribuyen a este futuro sombrío: su rápida difusión se produce sin que exista la posibilidad de implementar unos controles efectivos. Los gobiernos siguen mostrándose lentos a la hora de responder, ya sea por trabas legales, falta de inversión o por haber subestimado inicialmente la amenaza. Las élites políticas son incapaces de entender la magnitud de la brecha de seguridad que se está creando. Lo que los servicios de inteligencia pueden hacer también puede ser replicado por los grupos terroristas más capaces, especialmente si cuenta con algún tipo de respaldo por parte de regímenes políticos dispuestos a instrumentalizar el terrorismo para alcanzar sus objetivos.

Existen organizaciones terroristas que consiguen incorporar a sus filas especialistas en el empleo de la IA, reclutados de universidades punteras fuera del ámbito occidental, operadores desmovilizados de conflictos bélicos que han finalizado, o simplemente autoformados mediante los abundantes recursos disponibles en internet. Estos "tecnólogos del terror" generan capacidades que se adaptan de manera específica a las necesidades operativas de las organizaciones violentas.

La propaganda alcanza niveles de sofisticación sin precedentes. Los *deepfakes* son indistinguibles de contenido genuino incluso para los sistemas de detección

forense. Hay vídeos de ejecuciones fabricados completamente mediante IA circulando en la red, generando ansiedad social sin la necesidad de una violencia real. Cada usuario recibe mensajes de radicalización específicamente optimizados para sus características psicológicas, miedos, frustraciones y vulnerabilidades, las cuales han sido identificadas mediante el análisis automatizado de su huella digital.

Las operaciones terroristas se vuelven más letales y difíciles de prevenir. Los grupos utilizan análisis predictivo no solo para evitar su detección —tarea para la cual los Estados mantienen una cierta ventaja en datos— sino para optimizar el impacto de sus ataques. Los ataques de drones autónomos se vuelven rutinarios: enjambres de pequeños UAVs comerciales, coordinados mediante IA, atacan simultáneamente múltiples objetivos blandos, saturando las capacidades de respuesta policial. Pero también se disparan los ataques selectivos. Los terroristas se plantean la viabilidad de los asesinatos contra objetivos de alto impacto. Los algoritmos rastrean durante meses las apariciones públicas de las víctimas y sus redes sociales para predecir con precisión dónde puede ser encontrado, posteriormente, un dron con reconocimiento facial aguarda oculto en las inmediaciones y ejecuta con éxito el ataque en el momento adecuado. Este tipo de ataques quirúrgicos, pero automatizados, genera una profunda incertidumbre en las clases dirigentes. La sensación de vulnerabilidad erosiona la moral de los gobiernos y socava la confianza ciudadana.

La financiación terrorista se hace casi imposible de rastrear. Las criptomonedas, combinadas con sistemas de IA que optimizan constantemente las rutas de blanqueo, permiten flujos de fondos prácticamente indetectables. Los sistemas automatizados generan redes de empresas fantasma, cuentas intermedias y transacciones aparentemente legítimas que dificultan el análisis incluso para los algoritmos más sofisticados de inteligencia financiera.

En los regímenes autoritarios y en las zonas donde se da un vacío de autoridad, ciertos gobiernos operan de manera clandestina para proveer a los grupos terroristas de herramientas avanzadas para que actúen como *proxies*. La difuminación de los límites entre el terrorismo y la guerra híbrida entre Estados paraliza la acción colectiva internacional.

Los servicios de seguridad, aunque también están equipados con tecnologías punteras, se encuentran en una carrera tecnológica donde las ventajas son temporales y asimétricas. Cada sistema de detección se enfrenta a múltiples opciones de evasión. El terrorismo no aumenta necesariamente en frecuencia —los ataques exitosos pueden incluso disminuir— pero cuando ocurren son devastadoramente efectivos y traumáticamente impredecibles.

Los Estados, empujados por los acontecimientos, pueden llegar tomar medidas extremas y contraproducentes. Por desesperación, algunos gobiernos restringen internet o bloquean plataformas virtuales dañando las libertades y la economía digital. Otros adoptan vigilancia masiva intrusiva intentando cazar terroristas antes de que actúen, lo que genera controversia y erosiona las instituciones democráticas. En suma, el escenario pesimista no solo ve a terroristas triunfantes en golpes tácticos, sino produciendo un clima de crisis prolongada que lleva a errores estratégicos de los propios Estados, debilitando los valores que buscan proteger.

Escenario Intermedio: un equilibrio inestable

El escenario de equilibrio es el más complejo de describir. No es ni la victoria decisiva del Estado ni el triunfo de un terrorismo sobrealimentado por el algoritmo, sino un equilibrio dinámico e inestable donde ventajas y vulnerabilidades se distribuyen de manera desigual y cambiante. En este futuro, muchos de los hechos que se describen en los escenarios anteriores efectivamente ocurren, pero se compensan mutuamente.

La IA proporciona ventajas significativas al contraterrorismo en algunos dominios mientras crea vulnerabilidades explotables por sus adversarios en otros. El ritmo del cambio es tan rápido que surgen divisiones temporales claras: lo que hoy da ventaja al terrorista, en unos meses queda anulado, y viceversa. Los Estados democráticos mantienen supremacía en áreas que requieren recursos masivos: supercomputación, acceso a datos de grandes prestatarios de servicios de comunicación, coordinación internacional, pero se enfrentan a asimetrías en áreas donde los pequeños grupos pueden ser desproporcionadamente efectivos.

La propaganda es el ejemplo paradigmático. Las plataformas de internet mayoritarias desarrollan capacidades sofisticadas de detección y eliminación de contenido

extremista, pero el contenido simplemente migra a espacios menos regulados y reacios a cooperar: aplicaciones de mensajería encriptada, redes descentralizadas, plataformas de nicho con moderación laxa. El juego del gato y el ratón se acelera mediante IA en ambos lados, pero nunca se resuelve definitivamente.

Los grandes ataques terroristas disminuyen en frecuencia, pero aumentan en sofisticación. Los grupos más primitivos son efectivamente neutralizados por la vigilancia mejorada. Los atacantes solitarios son detectados en etapas cada vez más tempranas. Pero persiste un núcleo irreductible de células profesionales, técnicamente competentes y con una clara vocación hacia el largo plazo. Estos grupos, adoptan la IA selectivamente, allí donde proporciona una clara ventaja, ejecutan ataques espaciados en el tiempo, pero devastadores en impacto.

Una característica del escenario intermedio es que coexisten periodos de calma y picos de crisis. Por ejemplo, tras un año de ningún atentado importante (los estados han frustrado varios, la disuasión parece funcionar), un evento inesperado rompe la racha: un nuevo grupo con gran conocimiento tecnológico surge y comete un ataque exitoso que obliga a actualizar todos los protocolos. La sensación pública es de incertidumbre, pero no de pánico constante; la amenaza terrorista es solo una de las varias preocupaciones de seguridad, intermitente en su manifestación. La ciudadanía aprende a vivir con un riesgo manejable pero latente. La sociedad desarrolla mayor resiliencia psicológica al terrorismo, en parte porque la saturación de contenido sintético genera cierta inmunización. Cuando todo puede ser falso, la conmoción ante el contenido extremo disminuye. Los *deepfakes* se vuelven tan comunes que su impacto propagandístico se diluye. Sin embargo, esto erosiona la confianza en la información verificable, generando un entorno donde el debate público sobre las amenazas reales se contamina constantemente con todo tipo de desinformación.

No hay un tratado global sobre IA y terrorismo, aunque sí cooperación técnica en foros multilaterales para intercambiar mejores prácticas. Cada país lidia con el problema según sus capacidades: las naciones avanzadas tecnológicamente mantienen a raya la mayoría de las amenazas, mientras que las regiones más inestables sufren en mayor medida la brecha tecnológica. Esta desigualdad hace que persista la violencia política en zonas donde los Estados son más frágiles, donde el uso de la IA es más

rudimentario, pero suficiente para para desafiar a las fuerzas locales. En cambio, en los países más desarrollados, aunque siguen produciéndose intentos de atentados, la mayoría de las operaciones consiguen ser desbaratadas gracias a una infraestructura de seguridad inteligente.

Los dilemas éticos de la IA en contraterrorismo se vuelven batallas políticas permanentes. Los sistemas de vigilancia predictiva existen y funcionan, pero su uso está perpetuamente cuestionado, limitado por regulaciones en constante cambio, y sujeto a escrutinio judicial. Cada ataque exitoso genera demandas de mayor vigilancia. Cada falso positivo que arruina una vida, supone una enmienda al sistema. El péndulo oscila continuamente.

En este escenario, el terrorismo no desaparece ni se transforma en una amenaza existencial. Persiste como una amenaza manejable pero no eliminable, requiriendo vigilancia constante, adaptación y aceptación de que la seguridad absoluta es inalcanzable. La IA transforma las herramientas disponibles para ambos lados, pero no altera fundamentalmente el equilibrio estratégico del conflicto.

¿Qué nos dicen estos escenarios?

Los escenarios expuestos en este documento no son predicciones sino exploraciones de posibilidades, intentos de cartografiar un espacio de futuros potenciales que dependen de decisiones que aún no hemos tomado y desarrollos que aún no han ocurrido.
Algunas lecciones, sin embargo, emergen con cierta claridad.

En primer lugar, la IA no es una panacea, ni una fatalidad inevitable. No resolverá mágicamente el problema del terrorismo ni convertirá a los grupos terroristas en amenazas existenciales. Es una herramienta, y como toda tecnología, su impacto depende de cómo se use, quién la usa, y en qué contexto.

En segundo lugar, las asimetrías importan más que las capacidades absolutas. El terrorismo ha funcionado siempre mediante explotación de estas brechas: grupos pequeños que consiguen imponer costes desproporcionados a Estados mucho más poderosos. La IA crea nuevas asimetrías en ambas direcciones. Los países

mantienen ventajas en recursos computacionales, acceso a datos y capacidad de coordinación. Pero los actores no estatales pueden ser desproporcionadamente efectivos en áreas donde la agilidad, la creatividad y la disposición a asumir riesgos compensen sus limitados recursos.

La dimensión cognitiva del conflicto se hará aún más central. El terrorismo ha sido siempre, en palabras de Brian Jenkins, "teatro"[21]. La violencia física es instrumental; el objetivo es psicológico y político. La IA transforma radicalmente el escenario donde este drama se representa. En un mundo donde la fabricación de realidades alternativas es trivial desde el punto de vista técnico, la lucha por la verdad se vuelve tan crucial como la batalla por la seguridad física.

Los dilemas éticos no tienen respuestas técnicas. Las decisiones sobre cuánta privacidad merece la pena sacrificar a cambio de seguridad, cuánta vigilancia es aceptable en una democracia, qué errores algorítmicos son tolerables o quién debe rendir cuentas cuando los sistemas de IA fallan, son fundamentalmente políticas y morales, no técnicas. Siempre existirá la tentación de “tecnocratizar” estas decisiones —delegarlas a algoritmos o expertos. Sin embargo, son decisiones que definen profundamente el tipo de sociedad que queremos ser.

Pero, ante todo, debemos tener claro que la peor respuesta a la incertidumbre es la parálisis. Que el futuro sea incierto no significa que todas las opciones sean igualmente probables o deseables. Las decisiones que tomemos hoy —sobre gobierno de la IA, marcos regulatorios, inversiones en capacidades, educación pública, etc.— influirán significativamente en qué escenarios se materializan y cuáles no.

El horizonte 2035 no está tan lejos. Las tecnologías que definirán ese futuro ya existen y maduran a un ritmo frenético. Los grupos terroristas que operarán entonces ya existen. Las agencias de seguridad que los combatirán son las actuales. No estamos especulando sobre un futuro remoto, sino preparándonos para uno inminente.

[21] JENKINS, B. M. “International Terrorism: A New Mode of Conflict”, *Research Paper*, RAND Corporation, 1974.

La historia nos enseña que la tecnología raramente resuelve nuestros problemas fundamentales. Internet no eliminó el terrorismo; lo transformó. La IA hará lo mismo. La cuestión no es si el terrorismo persistirá, sino qué forma adoptará y qué precio estamos dispuestos a pagar para combatirlo.

*Manuel R. Torres Soriano**
Catedrático de Ciencia Política en la Universidad Pablo de Olavide de Sevilla
@mrtorsor