Documento
Opinión

61/2023 16 de junio de 2023

*Eduardo Rodríguez, Ramón Morales, Diego Crescente, María Isabel Cabello, Ignacio Paz**

La inteligencia artificial en la guerra híbrida como arma de desinformación

La inteligencia artificial en la guerra híbrida como arma de desinformación

Resumen:

La evolución de la tecnología está afectando a casi todos los ámbitos de la sociedad, pero también, y de manera muy significativa, al de la seguridad y la defensa. El conflicto de Ucrania ha demostrado que ahora las guerras se libran tanto en el espacio físico, como en el digital. Los ciberataques, las campañas de desinformación y las *fake news* son una amenaza para las sociedades democráticas.

La desinformación como arma de estrategia híbrida es tan antigua como la propia guerra, pero estamos asistiendo a un nuevo hito histórico, que está siendo definido como «cuarta revolución industrial», en el que la inteligencia artificial juega un papel determinante, que puede cambiar el escenario global. La soberanía tecnológica en IA debe de ser un objetivo estratégico, con la participación conjunta y coordinada de Administraciones públicas, empresas tecnológicas, sistemas educativos y agencias de verificación.

Palabras clave:

Inteligencia artificial, *fake news*, *deepfakes,* desinformación, infoxicación, guerra híbrida.

Cómo citar este documento:

RODRÍGUEZ LORENZO, Eduardo *et al. La inteligencia artificial en la guerra híbrida como arma de desinformación*. Documento de Opinión IEEE 61/2023. https://www.ieee.es/Galerias/fichero/docs_opinion/2023/DIEEEO61_2023_EDUROD_Inteligencia.pdf y/o enlace bie[3] (consultado día/mes/año)

La desinformación como arma en la guerra híbrida. ¿Un concepto nuevo?

Aunque el concepto intrínseco de la guerra híbrida es ancestral, su denominación se remonta a los últimos años del siglo XX, cuando, en un artículo del general James N. Mattis y del teniente coronel Frank G. Hoffman, en la prestigiosa revista *Naval Institute Proceedings* titulado «La guerra del futuro: el nacimiento del conflicto híbrido», se asientan las bases de esta teoría y la denominación adquiere carta de naturaleza.

Según esta doctrina, en la guerra híbrida se utilizan toda clase de medios al alcance de los contendientes, no solo mediante el empleo de la fuerza convencional, sino apoyándose en cualquier otro medio, como el terrorismo, la insurgencia, la migración, los recursos naturales, o *las técnicas de influencia sobre la población***.**

El general Sun Tzu, en su libro *El arte de la guerra* escrito hace más de 2.500 años en la antigua China, ya revela los principios del alto valor estratégico de la desinformación como arma incuestionable. Conceptos como «saber practicar el arte del engaño para confundir y debilitar al enemigo», o «desacredita cuanto está bien en el país del adversario», resultarán vitales en multitud de conflictos, tanto para influir en el enemigo, como en la percepción y sentimientos de la población propia.

El comité para «agitprop», dependiente del partido comunista, aplicó un novedoso método de propaganda durante la revolución rusa. Se preparaban trenes equipados con sistemas de proyección e imprentas, llenos de miembros del partido y se enviaban a poblaciones lejanas para agitar y organizar a la población.

Ilustración 1. Trenes AGITPROP (1920).
https://mediartinnovation.com/

La historia de los conflictos armados y de las relaciones internacionales han demostrado que, a través del engaño, afectando directamente a la naturaleza humana y sus percepciones, se pueden conseguir grandes victorias sin necesidad de emplear la fuerza.

Mediante la desinformación se puede llegar a *socavar la moral de un ejército enemigo o, lo que es más importante, de su población,* con el fin de influir en sus gobernantes para que cesen en su actitud beligerante, o conseguir ventajas en una potencial negociación.

La información sobre las propias fuerzas, fuerzas aliadas y sobre las fuerzas enemigas ha sido siempre un factor clave en las operaciones militares. En palabras de Sun Tzu: «Si te conoces a ti mismo y conoces a tu enemigo, no necesitas temer al resultado de un centenar de batallas. Si te conoces a ti mismo, pero no conoces a tu enemigo, por cada victoria que ganes sufrirás también una derrota. Si no te conoces ni a ti mismo ni a tu enemigo, sucumbirás en cada batalla».

Durante el asalto final al continente europeo en la Segunda Guerra Mundial se llevó a cabo una de las operaciones de engaño a gran escala más relevantes de la historia. El objetivo era confundir a los nazis sobre lugar y fecha del desembarco aliado. La operación involucró el envío de desinformación por radio, el despliegue de carros de combate «hinchables» y la filtración de planes falsos.

Ilustración 2. Operation Bodyguard (1943-1944).
https://www.historyhit.com/

La guerra de la información es por tanto tan antigua como la guerra misma. Es una forma de combatir, extensamente empleada en conflictos que se mueven en la denominada «zona gris», donde además de su bajo coste relativo, tiene como ventaja añadida su difícil atribución.

Las *fake news*

La propagación de información errónea es un tema crítico que preocupa a los Estados democráticos. Las noticias falsas, también conocidas como *fake news*, pueden ocultar intenciones maliciosas destinadas a manipular la opinión pública y debilitar la estabilidad de las instituciones.

Si hay un genio de las noticias falsas o fake news («yellow news» por aquel entonces), es, sin duda, el magnate William Randolph Hearst. A finales del siglo XIX, España y Estados Unidos se enfrentaban por el dominio de Cuba. Hearst ordenó a sus periodistas (era dueño de más de 20 periódicos) que exageraran las noticias sobre la insurrección cubana, hasta que esto culminó en la explosión de El Maine en 1898. Aunque las causas del suceso nunca estuvieron claras, el magnate apuntó con sus titulares al enemigo, lo que provocó irremediablemente la guerra y que nuestro país perdiera la colonia.

Ilustración 3. Explosión del Maine. Age fotostock

La difusión de bulos y la desinformación plantean una amenaza global para la libertad y la democracia, especialmente en la era digital donde la propagación de estas campañas se realiza a una velocidad alarmante. En los últimos años, tanto los flujos de información como los de desinformación han aumentado en similar medida, lo que hace que este problema sea acuciante.

La Comisión Europea ha identificado que la coordinación de tres grupos de actores es fundamental en la lucha contra la desinformación. Estos grupos son las empresas tecnológicas, la sociedad civil, que incluye a los verificadores de información (*fact checkers*), y a las instituciones académicas.

La comunicación estratégica es una herramienta clave en la lucha contra la desinformación y requiere un enfoque integral. Las campañas de desinformación suelen ser más que la mera difusión de noticias falsas, ya que tienen como objetivo construir un relato malintencionado. Por esta razón, la lucha contra la desinformación se encuentra estrechamente ligada a la comunicación estratégica, la diplomacia pública y la comunicación digital.

España ha asumido un fuerte compromiso en la lucha contra la desinformación y ha estado trabajando activamente en colaboración con la Unión Europea, con especial énfasis desde el año 2018, para establecer procedimientos ágiles y efectivos para abordar este problema. Asimismo, la OTAN ha asumido un papel activo en la lucha contra la desinformación, que se ha convertido en elemento fundamental de su estrategia de comunicación en el contexto de las nuevas amenazas híbridas.

La inteligencia artificial como arma de desinformación

La inteligencia artificial (IA) y el aprendizaje automático se han convertido en un *elemento tecnológico disruptivo* que está cambiando las reglas del juego en muchos ámbitos de nuestra sociedad, siendo de gran ayuda para acelerar procesos manuales, reducir costes y eliminar errores humanos. Sin embargo, también pueden ser empleadas como *arma para realizar ataques de desinformación mucho más eficientes***.**

Un ejemplo de amenaza impulsada por la inteligencia artificial son los *deepfakes*, una combinación de *deep learning* (aprendizaje profundo) y *fake* (falso). Se utilizan para manipular rostros y voces, imitando gestos y patrones de habla para crear un engaño que, con el estado del arte actual, hace casi imposible diferenciar lo que es real y lo que es falso. Aunque estas tecnologías se popularizaron inicialmente como «diversión», también pueden ser empleadas, por parte de actores hostiles, como potentísima arma de desinformación en la guerra híbrida.

Ilustración 4. *Deepfake* del presidente Zelensky publicado en Facebook y Youtube el 16/3/2022, pidiendo la rendición de sus tropas ante el ejército ruso.

El exponencial avance de las tecnologías de inteligencia artificial, y el empleo simultáneo de diferentes capacidades de IA, como los *deepfakes*, la generación automática de avatares en redes sociales y de contenidos, se convierten en un factor multiplicador de la amenaza, que tiene como objetivo influir en la opinión pública y los medios de

comunicación, y que resalta la importancia de adoptar una serie de medidas que necesariamente también deben apoyarse en estas nuevas tecnologías.

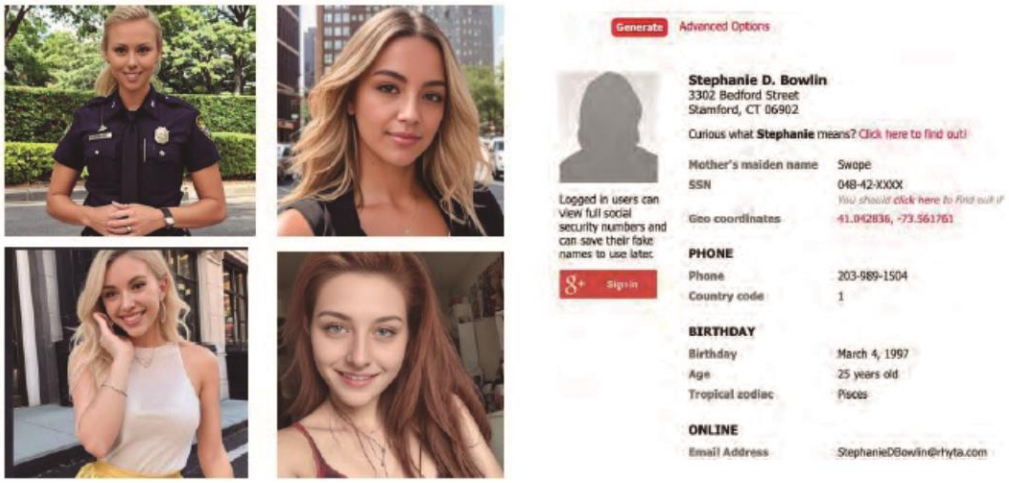

Ilustración 5. Avatares generados por inteligencia artificial

Por otra parte, los modelos de inteligencia artificial y aprendizaje automático requieren cantidades masivas de datos de entrenamiento para poder funcionar eficientemente. Cuantos más datos se introduzcan en el sistema, más preciso será. Sin embargo, también es posible engañar a estos sistemas para que cometan errores, alimentándolos con datos de entrenamiento incorrectos o mediante el envenenamiento de datos o *infoxicación*.

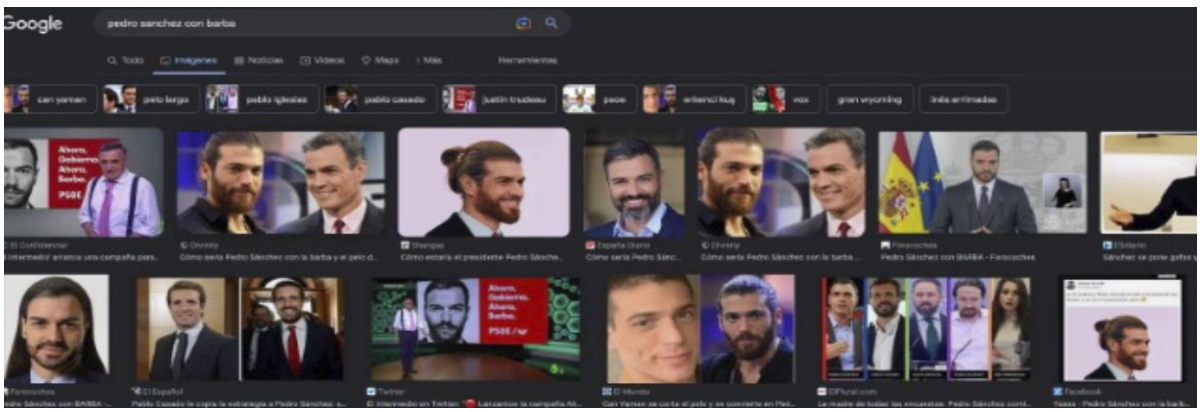

Ilustración 6. A la petición «Describe físicamente a Pedro Sánchez», ChatGPT responde, entre otras características que, «en cuanto a su rostro, *suele lucir una barba cuidada y recortada* que le da un aspecto más maduro y distinguido». Este error es producto de la «infoxicación» involuntaria provocada en enero de 2020 por una campaña en redes sociales promovida, en tono jocoso, por un conocido programa de TV.

En el ámbito de la ciberdefensa, un ataque de *botnet* impulsado por la IA también puede ser extremadamente peligroso. Un atacante puede recopilar todos los datos posibles y

entrenar al *malware* para que pueda predecir el patrón de defensa de un sistema y cambiar automáticamente su estrategia de ataque en función de la respuesta del sistema de defensa. Este tipo de ataques podría extenderse fácilmente por múltiples dispositivos y redes, con lo que caerían las infraestructuras críticas de un Estado en un momento conveniente para el atacante.

La inteligencia artificial, también como herramienta contra la desinformación

La guerra de Ucrania ha acelerado el empleo de la IA como arma, pero también como herramienta de defensa. Durante el evento «Deconstruyendo la desinformación» en el marco del Mobile World Congress 2022, se destacó el ejemplo de la «guerra híbrida», al detectar más de 1.600 noticias falsas en las tres primeras semanas después de la invasión de Ucrania. Desde el Gobierno ruso se comenzaron a emplear cuentas institucionales en medios afines y en redes sociales, para difundir mensajes falsos en apoyo de la versión del Kremlin, según los estudios realizados por *fact-checkers* u organismos de verificación acreditados.

Esto refuerza la importancia de desarrollar herramientas para detectar este tipo de acciones, y es donde la IA también puede ser aplicada como elemento fundamental.

Mediante una combinación de inteligencia artificial y modelos estadísticos, una plataforma de IA puede analizar el vocabulario, la semántica y los patrones de lenguaje, y determinar la probabilidad de que un texto haya sido generado artificialmente. Para ello, es necesario encadenar diferentes tecnologías en continuo desarrollo, como el NLP (procesamiento de lenguaje natural), el STS (similitud semántica textual), el NLI (motor de inferencia de lenguaje natural), la minería de datos para detección de anomalías, y el XAI (inteligencia artificial explicable).

Asimismo, mediante el empleo de algoritmos de IA programados para ello y *en permanente adaptación*, se pueden detectar contenidos manipulados o creados artificialmente en imágenes, vídeos, audios, perfiles de redes sociales, etc.

Este tipo de sistemas, en conexión con los recursos proporcionados por organizaciones como la International Fact-Checking Network, creada en 2015 para reunir a la creciente

comunidad de verificadores de datos de todo el mundo, pueden inferir si un texto publicado en cualquier medio o red social es o no una noticia falsa.

La IA al servicio del ciberespionaje: Katie Jones, logró que personajes muy influyentes de la vida política de Washington aceptaran su solicitud de conectar en la red LinkedIn.
Mediante herramientas de inteligencia artificial, se identificó como perfil falso, *que había sido generado con técnicas Deep Fake. Con alta probabilidad, era controlado por un servicio de inteligencia hostil, como agente Ciberhumint.*

Análisis FODA

En el ámbito de la sociedad de la información en la que el mundo se encuentra ya inmerso, queda patente que la inteligencia artificial (IA) se constituye como una herramienta muy potente, que puede mejorar significativamente la forma en que procesamos, almacenamos y utilizamos la información. Sin embargo, también presenta algunas amenazas significativas, por lo que resulta necesario realizar un análisis de riesgos más detallado, con la finalidad de aprovechar las fortalezas y oportunidades, minimizando las amenazas y las debilidades.

Una de las herramientas para efectuar este análisis es la *matriz FODA*, en la que se exponen las *fortalezas, oportunidades, debilidades, y amenazas*, para crear un diagnóstico, o punto de situación. Este análisis facilitará el diseño posterior de las estrategias y las acciones a implementar.

A continuación, se realiza un diagnóstico de cada uno de los parámetros señalados:

Fortalezas

- **Capacidad de procesar grandes cantidades de datos**. La IA puede procesar grandes cantidades de información en tiempo real, lo que puede ayudar a identificar patrones, tendencias y anomalías.
- **Automatización de tareas repetitivas**. La IA puede automatizar tareas repetitivas y monótonas, lo que puede liberar tiempo y recursos de las personas, que pueden enfocarse en tareas más importantes y estratégicas.
- **Mejora de la precisión y la exactitud**. La IA puede mejorar la precisión y la exactitud de los resultados en comparación con los métodos tradicionales de procesamiento de información.
- **Capacidad de aprendizaje automático**. La IA puede aprender y mejorar a partir de los datos que se le proporcionan, lo que puede ayudar a mejorar la calidad y relevancia de los resultados.
- **Identificación de patrones y tendencias**. La IA puede identificar patrones y tendencias en grandes conjuntos de datos, que pueden no ser evidentes, lo que puede ayudar a identificar tanto oportunidades como problemas.
- **Análisis predictivo**. La IA puede ser utilizada para realizar análisis predictivos, basados en hechos o tendencias pasadas, lo que puede ayudar a la toma de decisiones.

Oportunidades

- **Mayor eficiencia y productividad**. La automatización de tareas, que puede proporcionar la IA, puede permitir la redistribución de recursos, tanto humanos como materiales, para optimizar los procesos productivos, agilizar trámites administrativos, e incrementar la eficiencia de empresas y organizaciones.
- **Identificación de oportunidades de negocio**. El análisis automático de patrones y tendencias, sobre grandes conjuntos de datos, puede ayudar a identificar oportunidades de negocio y áreas de mejora.

- **Descubrimiento de conocimientos**. Mediante el análisis asistido por IA de grandes conjuntos de datos se puede llegar a conclusiones que no son evidentes para un análisis humano, lo que puede conducir a innovaciones tecnológicas y científicas en muy variadas áreas del conocimiento.
- **Mejora de la seguridad de la información**. La IA puede ayudar a mejorar la seguridad de la información al detectar patrones de comportamiento sospechosos y amenazas potenciales. Esto puede incluir la detección de fraude en línea, la identificación de patrones de actividad sospechosa y la predicción de posibles vulnerabilidades en sistemas y redes de información.
- **Personalización de experiencias.** La IA puede ser utilizada para personalizar la experiencia del usuario en el uso de servicios y aplicaciones.

Debilidades

- **Falta de transparencia**. La IA puede ser difícil de entender y explicar debido a la complejidad de los algoritmos usados. Esto puede dificultar la evaluación y la validación de los resultados, lo que puede afectar a la confianza en las herramientas basadas en IA.
- **Dependencia de los datos**. La IA depende de datos precisos y representativos para producir resultados rigurosos y relevantes. Si los datos están incompletos o sesgados, los resultados también pueden ser inexactos o incompletos.
- **Inversión elevada**. La implementación de la IA puede ser costosa, ya que se requiere *hardware* muy potente y *software* especializado, así como personal capacitado para diseñar, entrenar y mantener algoritmos y modelos de datos.
- **Aspectos éticos**. La IA puede ser empleada de manera no ética o ilegal por parte de actores hostiles y Estados no democráticos, lo que puede poner en peligro los derechos individuales de los ciudadanos.
- **Falta de creatividad y comprensión contextual**. Aunque la IA es muy efectiva en el procesamiento y análisis de grandes cantidades de información, puede tener dificultades para comprender el contexto y la creatividad que a menudo se requiere en la toma de decisiones y el pensamiento estratégico.

- **Falta de comprensión humana**. La IA puede tener dificultades para comprender las sutilezas de la información humana, como el lenguaje natural y las emociones, lo que puede llevar a malentendidos o interpretaciones erróneas.

Amenazas

- **Manipulación de la información**. La IA puede ser utilizada para manipular información mediante la generación de noticias falsas o sesgadas y la creación de perfiles falsos en redes sociales que las propaguen. Esto puede tener graves consecuencias para la sociedad, ya que la información errónea puede influir en la opinión pública y, en consecuencia, a la toma de decisiones.
- **Sesgo algorítmico**. La IA aprende a partir de los datos que se le proporcionan, y si los datos están sesgados, los resultados que proporciona también lo estarán. Estos sesgos pueden estar provocados, bien por motivos pasivos, como en el caso de disponer de repositorios de información incompletos, o por causas activas, en el caso del empleo de técnicas de «infoxicación» o manipulación por parte de actores hostiles.
- **Robo de datos y privacidad**. La IA puede usarse para recopilar grandes cantidades de datos personales, lo que puede representar una amenaza para la privacidad de la ciudadanía de un estado objetivo. Si estos datos caen en manos de potenciales adversarios, pueden utilizarse para campañas de desinformación, búsqueda de vulnerabilidades u otros tipos de ataques, que podrían, además, realizarse de manera automática y masiva.
- **Dependencia tecnológica.** La IA, junto con tecnologías como el 5G o el Internet de las cosas (IoT) se están convirtiendo en un motor de cambio, cada vez más relevante. La dependencia exclusiva en estas tecnologías, *si no se adoptan las medidas de seguridad adecuadas*, pueden tener consecuencias catastróficas, como en el caso de ataques automatizados a sistemas de gestión y control, especialmente sobre infraestructuras críticas.
- **Crisis de opinión**. Por otra parte, la dependencia excesiva en la IA puede llevar a una falta de pensamiento crítico y a una reducción de la capacidad de tomar

decisiones independientes, sobre todo, si como hemos visto, el conjunto de datos sobre los que trabaja ha podido ser manipulado mediante técnicas de infoxicación, o mediante ataques de desinformación.

Ilustración 8. Matriz FODA: Inteligencia artificial en el ámbito de la información

Estrategias y acciones nacionales en el ámbito de la inteligencia artificial

España parte de una posición favorable para abordar esta revolución tecnológica y científica, contando con infraestructuras de calidad para poder desarrollar actividades relacionadas con la inteligencia artificial (IA).

No obstante, este proceso también plantea importantes retos, por ejemplo:

- Aumentar las competencias digitales de la población, en especial la de las personas en situación o riesgo de exclusión social.
- Acelerar la digitalización del tejido de pequeñas y medianas empresas (PYMES).
- Promover la creación de repositorios de datos y facilitar el acceso a los mismos.

- Mejorar la eficiencia y productividad de los servicios públicos.
- Estimular la colaboración para incrementar la inversión pública y privada en I+D+i.

Para dar respuesta a estos retos, se han adoptado las siguientes iniciativas a nivel nacional, que se desarrollan en los siguientes apartados:

Estrategia Nacional de Inteligencia artificial (ENIA)

La Estrategia Nacional de Inteligencia Artificial tiene como objetivo proporcionar un marco de referencia para el desarrollo de una IA inclusiva, sostenible y centrada en la ciudadanía. La ENIA es uno de los ejes de la Agenda España Digital 2026 y uno de los componentes del Plan de Recuperación, Transformación y Resiliencia de la economía española, vertebrando la acción de las distintas administraciones y proporcionando un marco de referencia e impulso para el sector público y privado en este ámbito. La elaboración de la Estrategia Nacional responde al compromiso compartido con nuestros socios europeos, para que la UE se sitúe como líder en esta materia.

Objetivos de la Estrategia

- Excelencia científica e innovación en IA.
- Proyección de la lengua española.
- Creación de empleo cualificado.
- Transformación del tejido productivo.
- Entorno de confianza con relación a la IA.
- Valores humanistas en la IA.
- IA inclusiva y sostenible.

Plan de acción

Para dar cumplimiento a los objetivos de la Estrategia se han definido seis ejes de actuación, que agrupan a las acciones prioritarias que se llevarán a cabo a lo largo del periodo 2020-2025.

- **Eje estratégico 1**. Impulsar la investigación científica, el desarrollo tecnológico y la innovación en IA.

- **Eje estratégico 2**. Promover el desarrollo de capacidades digitales, potenciar el talento nacional y atraer talento global.
- **Eje estratégico 3**. Desarrollar plataformas de datos e infraestructuras tecnológicas que den soporte a la IA.
- **Eje estratégico 4**. Integrar la IA en las cadenas de valor para transformar el tejido económico.
- **Eje estratégico 5**. Potenciar el uso de la IA en la Administración pública y en las misiones estratégicas nacionales.
- **Eje estratégico 6**. Establecer un marco ético y normativo que refuerce la protección de los derechos individuales y colectivos, a efectos de garantizar la inclusión y el bienestar social.

Esta Estrategia, además, es instrumental a la hora de afrontar grandes desafíos sociales para una IA inclusiva y sostenible, por ejemplo, para reducir las brechas de género y digital y para favorecer la vertebración territorial y la transición ecológica.

La puesta en marcha de esta Estrategia requiere la movilización de un importante volumen de inversión, tanto pública como privada. La inversión pública, por parte del Estado, asciende a un total de 600 millones de euros en el periodo 2021-2023, a los que se añadiría el fondo «Next Tech» para impulsar el emprendimiento en tecnologías digitales habilitadoras. Este impulso, que podría movilizar una inversión privada de unos 3.300 millones de euros, servirá de catalizador de la acción de las universidades y empresas, orientando las prioridades, generando sinergias y cubriendo aquellas áreas en que el mayor riesgo o la falta de mercados puede suponer un retraso para la iniciativa privada.

Creación de la Agencia Española de Supervisión de la IA (AESIA)

La AESIA es la futura agencia estatal, con sede en La Coruña, cuya creación fue anunciada por el Gobierno de España el 28 de diciembre de 2021 y que se encargará de supervisar el cumplimiento de la regulación europea en materia de IA.

La AESIA está adscrita a la Secretaría de Estado de Digitalización e Inteligencia Artificial dentro del Ministerio de Asuntos Económicos y Transformación Digital, con una dotación presupuestaria de 5 millones de euros.

Esta nueva Agencia Estatal se configura como «ente con personalidad jurídica pública, con patrimonio propio y autonomía en su gestión y potestad administrativa, que actuará con plena independencia orgánica y funcional de las Administraciones públicas, de forma objetiva, transparente e imparcial, llevando a cabo medidas destinadas a la minimización de riesgos significativos sobre la seguridad y salud de las personas, así como sobre sus derechos fundamentales, que puedan derivarse del uso de sistemas de IA».

Objetivos y competencias

Entre sus objetivos se encuentran los siguientes: promover el desarrollo y uso responsable, sostenible y confiable de la inteligencia artificial; la concienciación, divulgación y promoción de la formación; la definición de mecanismos de asesoramiento y atención; la colaboración y coordinación con otras autoridades, nacionales y supranacionales, de supervisión de la IA; el fomento de entornos reales de prueba de los sistemas de IA, para reforzar la protección de los usuarios y minimizar los riesgos que puede traer la IA en campos como la seguridad, la intimidad y la salud de las personas, así como sobre los demás derechos fundamentales.

Quedan en el aire varias cuestiones, entre otras, si tendrá capacidad sancionadora, o el ámbito competencial para evitar conflictos con otros organismos, como la Agencia Española de Protección de Datos, el Instituto Nacional de Ciberseguridad, la Comisión Nacional de Mercados de la Competencia o la Secretaría de Estado de Digitalización e IA (SEDIA).

España se convertirá en el primer país de la Unión Europea con una agencia estatal de supervisión de la IA, anteponiéndose a la entrada en vigor del futuro reglamento aplicable a los sistemas de IA (previsto para principios de 2024), que establecerá la necesidad de que los Estados miembros cuenten con una autoridad supervisora en esta materia.

Conclusiones

La imparable y vertiginosa evolución de la tecnología ha transformado la forma en la que vivimos, trabajamos, nos comunicamos y educamos, y está afectando fuertemente a casi todos los ámbitos de la sociedad, *pero también, y de manera muy significativa, al de la seguridad y la defensa.*

El conflicto de Ucrania ha demostrado que ahora las guerras se libran tanto en el espacio físico, como en el digital, en un teatro a escala global. Los ciberataques, las campañas de desinformación y las *fake news* son una amenaza para la seguridad y la estabilidad de las sociedades democráticas, y constituyen además una potente arma de control o influencia sobre su población.

La tecnología permite que cualquier persona con conexión a Internet pueda estar informada sin límites ni barreras territoriales, y, sin embargo, a pesar de tener un mayor acceso a contenidos, noticias y fuentes directas, puede crearse la paradoja de una sociedad más desinformada que nunca en la historia de la humanidad.

Aunque la desinformación como arma de estrategia híbrida es tan antigua como la propia guerra, estamos asistiendo a un nuevo hito histórico, que está siendo definido como «cuarta revolución industrial», en el que la inteligencia artificial juega un papel determinante, que puede cambiar el escenario global.

La matriz FODA presentada es autoexplicativa, tanto en lo que a aspectos positivos se refiere (entendidos como la suma de fortalezas y oportunidades, que es necesario explotar), como y quizás hasta con mayor rotundidad, en lo relativo a los aspectos negativos (suma de debilidades y amenazas, que es preciso combatir).

La dependencia de datos fidedignos (de calidad, únicos y seguros), además de las fisuras existentes en los aspectos éticos del uso de la IA son quizás las debilidades más notorias y las que jugarán un papel más determinante en el corto y medio plazo.

Dentro de las debilidades tampoco podemos ser ajenos a la elevada inversión que el campo de IA requiere, ya que en ocasiones los presupuestos necesarios no son fáciles de argumentar a la sociedad, e incluso en ocasiones, no son bien entendidos y respaldados en el ámbito parlamentario.

En lo que respecta a las amenazas, desde la perspectiva de este estudio, la manipulación de la información es crucial, ya que puede llegar a ser determinante y así generar una corriente de opinión interesada, desdibujada y velada respecto a la realidad. Esta corriente interesada puede manipular a la sociedad y poner en riesgo la estabilidad de nuestras estructuras de gobierno, que son la base de la seguridad y, en consecuencia, los cimientos de nuestro estado de bienestar y prosperidad.

Además, para los regímenes democráticos no es fácil el demostrar la manipulación deliberada por terceros de la información y, además, una acusación poco fundamentada puede comprometer las relaciones entre Estados.

Por otra parte, el mantener una cierta soberanía tecnológica en IA, al menos el evitar una dependencia excesiva de otros países, fundamentalmente de aquellos de fuera del ámbito de la UE, debe de ser un objetivo estratégico, pues de lo contrario la vulnerabilidad en todos los aspectos clave podría ser determinante.

Por todo ello, resulta manifiestamente esencial la participación conjunta y coordinada de administraciones públicas a nivel nacional e internacional, empresas tecnológicas, sistemas educativos y agencias de verificación.

En el ámbito nacional, la Estrategia de Seguridad Nacional, la Estrategia Nacional de Inteligencia Artificial y, de manera muy notable y oportuna, la creación de la nueva Agencia Española de Supervisión de la IA (AESIA), resultan de capital importancia, no solo para aplicar la IA como un elemento positivo de transformación de la economía y la sociedad, sino como elementos imprescindibles para *observar, prevenir y combatir los riesgos y amenazas* que la irrupción de la inteligencia artificial está inyectando, de manera exponencialmente progresiva, en el escenario global.