

# Agencia Española de Supervisión de la Inteligencia Artificial, la clave para un desarrollo tecnológico ético, justo y sostenible

## ALEJANDRA ARTERO MUÑOZ

Experta en Comunicación y Directora de Comunicación en la Secretaría de Estado de Digitalización e Inteligencia Artificial.

## CARLOS FRANCISCO RUIZ DE TOLEDO RODRÍGUEZ

Letrado de la Administración de la Seguridad Social, responsable del Proyecto AESIA y vocal en dicho Consejo Rector y Subdirector General de Ayudas en la Secretaría de Estado de Digitalización e Inteligencia Artificial.

## PILAR MAIRAL MEDINA

Doctora en Ciencia Política, profesora Asociada del Departamento de Ciencia Política y Administración pública de la Universidad Complutense de Madrid y Asesora en la Secretaría de Estado de Digitalización e Inteligencia Artificial.

### RESUMEN

La AESIA es la Agencia Española de Supervisión de la Inteligencia Artificial, creada en 2023 para regular y supervisar el uso y desarrollo de los sistemas de IA en España. La AESIA tiene su sede en A Coruña, en el edificio La Terraza.

La AESIA se enmarca dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), que es una de las medidas de la Agenda España Digital 2026 y del Plan de Recuperación, Transformación y Resiliencia. La ENIA tiene como objetivo proporcionar un marco de referencia para el desarrollo de una IA inclusiva, sostenible y centrada en la ciudadanía, y mejorar la preparación y la competitividad del tejido productivo español en el ámbito de la IA.

Como paso previo al desempeño de la función supervisora de la inteligencia artificial, la AESIA prevé convertirse en un centro de referencia y análisis sobre la IA, que aportase una visión estratégica, crítica y constructiva sobre esta tecnología, que contribuyese al desarrollo de una IA responsable, inclusiva y sostenible, y que impulsase el liderazgo y la competitividad de España en este campo.

### PALABRAS CLAVE

Inteligencia artificial  
Estrategia Nacional de Inteligencia Artificial (ENIA)  
Think tank  
Normativa europea  
Ética  
Gobernanza

### ABSTRACT

*AESIA is the Spanish Agency for the Supervision of Artificial Intelligence, created in 2023 to regulate and supervise the use and development of AI systems in Spain. AESIA is based in A Coruña, in the La Terraza building.*

*AESIA is part of the National Artificial Intelligence Strategy (ENIA), which is one of the measures of the Spain Digital Agenda 2026 and the Recovery, Transformation and Resilience Plan. The ENIA aims to provide a framework for the development of an inclusive, sustainable and citizen-centered AI, and to improve the preparedness and competitiveness of the Spanish productive fabric in the field of AI.*

*As a previous step to the performance of the supervisory function of artificial intelligence, the AESIA envisages becoming a centre of reference and analysis on AI, providing a strategic, critical and constructive vision on this technology, contributing to the development of responsible, inclusive and sustainable AI, and boosting Spain's leadership and competitiveness in this field.*

### KEYWORDS

Artificial intelligence  
National Artificial Intelligence Strategy (ENIA)  
Think tank  
Ethics  
Governance

## Contexto

El avance de las tecnologías disruptivas y su impacto transversal, tanto a nivel económico como cultural, social y político, es un fenómeno incuestionable a nivel global. Durante los últimos años, estamos experimentando un proceso de transformación estructural que involucra e interpela a todos los gobiernos y las administraciones públicas a lo largo y ancho del mundo, en base a las oportunidades del avance digital, pero también a las incertidumbres, riesgos y amenazas que este supone. La historia nos ha mostrado que toda gran revolución de calado debe asentarse sobre una base sólida de principios y valores compartidos que nos permitan avanzar partiendo de un suelo firme sobre el que cultivar un progreso económico y social para la mayoría. Desde este punto de vista, un movimiento tectónico de la trascendencia que implica un desarrollo tecnológico sin precedentes como el que vivimos requiere garantizar unos cimientos sólidos que no permitan que la convivencia, el orden democrático y los derechos fundamentales de las personas se resquebrajen.

En este contexto, las políticas de transformación digital deben incorporar como objetivo esencial que tanto la ciudadanía como, las organizaciones privadas y las instituciones públicas cuenten con las herramientas y el contexto ético, cultural e intelectual adecuado de cara a evolucionar a la misma velocidad que esta innovación, cuidando en el proceso los derechos y las garantías que existen en lo que se conoce como mundo *offline*.

Los últimos acontecimientos en el avance y la integración social de tecnologías con un potencial disruptivo como la Inteligencia Artificial (IA) han demostrado que solo quien es capaz de generar certidumbre, seguridad y de garantizar un desarrollo sostenible y ético de la misma contará con el respaldo de la ciudadanía y de una comunidad internacional que ha abierto los ojos con respecto a la trascendencia histórica del escenario que se está conformando.

El avance de una digitalización humanista, llamada a colocar a las personas y sus derechos en el centro, ha puesto a Europa y España en una posición de privilegio al potenciar a nivel global esta visión humanística, al margen de las tendencias y corrientes coyunturales. Ese trabajo prospectivo y estratégico, basado en una mirada larga y estructural sobre el impacto de la IA, es el que ahora nos sitúa a la vanguardia del desarrollo de una IA ética, regulada y transparente que ya forma parte de un consenso global.

Este es el ámbito en el que España puede y debe potenciar su competitividad global. La inversión en innovación, infraestructuras, talento y en la integración de la IA en nuestras cadenas de valor se plantea necesariamente acompañada de un recorrido humanista ambicioso y a la vanguardia mundial. Una visión del impulso de la nuestra competitividad en base a principios y fortalezca nuestra posición en el mercado global. El desarrollo de una IA ética y fundamentada en los valores democráticos y humanistas europeos es el factor diferencial que debe atravesar nuestra propuesta y nuestro papel en una revolución que sin duda es tecnológica, pero también sustancialmente cultural, social e intelectual. Esta es la senda a través de la que Europa tiene la oportunidad de poner su sello, de impulsar su valor, de marcar un antes y un después. Ya hemos demostrado ir un paso por delante, y es el momento de hacer de ello nuestra fortaleza, nuestra esencia como bloque.

Hasta hace muy poco, Europa se encontraba sola defendiendo esta visión humanista como la única válida y con recorrido ante la irrupción de la IA, mientras que otros países usaban la desregulación y la falta de estructuras garantistas en torno a la innovación como una ventaja competitiva que el tiempo ha demostrado no ser solvente ni competitiva a medio y largo plazo.

Hoy nadie duda que el avance ético de la IA es una condición *sine qua non* para un desarrollo tecnológico a la altura de nuestras democracias y los principios y derechos fundamentales que las sustentan. El propio CEO de Open IA, empresa estadounidense desarrolladora del fenómeno de Chat GPT tras la que se encuentra el gigante tecnológico Microsoft, se manifestó al respecto afirmando la necesidad de una mayor regulación, argumentando la creación de organismos de supervisión de IA, tanto a nivel internacional como nacional. Hoy podemos decir que el consenso global en torno a esta visión ya es global.

Evolucionar y aumentar las capacidades propias del ser humano de la manera en la que la IA propone, replicar una inteligencia genuinamente humana para potenciarla en favor del progreso, requiere tener muy claro que las personas y el blindaje de sus derechos deben estar por encima de cualquier otro criterio. Para ello, es esencial, ante todo, partir de un conocimiento en profundidad de la condición humana como el que Europa puede ofrecer y propone como camino a seguir.

En el caso concreto de España, esta visión sobre la transformación digital es prioritaria. Así lo refleja la Agenda España Digital 2026, la hoja de ruta en materia digital, que marca como focos de esta digitalización de la economía y las empresas, las administraciones públicas y, por supuesto, las personas. Dicha Estrategia incluye diferentes planes estratégicos sectoriales entre los que se encuentra la Estrategia Nacional de Inteligencia Artificial (ENIA), que *«tiene como objetivo proporcionar un marco de referencia para el desarrollo de una IA inclusiva, sostenible y centrada en la ciudadanía»* (ENIA, 2021).

Una de las principales medidas recogidas en la Estrategia es la puesta en marcha de la Agencia Española de Supervisión de la Inteligencia Artificial (AESIA). Con su creación, España se convierte en el primer país europeo en tener un órgano de estas características y se anticipa a la entrada en vigor del Reglamento Europeo de Inteligencia Artificial.

Las competencias de este órgano comparten el fundamento ético anteriormente introducido: una visión humanista de la tecnología, que pretende poner a las personas en el centro de la transformación digital. En este sentido, España ha tomado este espíritu como marca diferencial en su transformación digital y, así, las iniciativas vinculadas a regulación, reformas, concienciación y sensibilización persiguen una misma meta: la garantía de los derechos, la ética y la sostenibilidad en el mundo online.

## La Estrategia Nacional de Inteligencia Artificial (ENIA): una política pública transversal con una visión transformadora de la economía

La política pública española de impulso de la Inteligencia Artificial establece desde 2020 la firme intención de impulsar la Inteligencia Artificial como elemento transversal de la transformación del tejido productivo. Tras el importante desarrollo de estas tecnologías en la última década, la IA ha demostrado ser un campo con un potencial sin precedentes de cara a avanzar hacia modelos industriales y de producción más eficientes, resilientes y sostenibles.

Un cambio de paradigma que ofrece a España una oportunidad única para abordar uno de los grandes problemas que durante décadas ha arrastrado nuestra economía, como es el de la falta de productividad y, por ende, de competitividad en el mercado global. La IA es una respuesta clave a todos los retos de competitividad y optimización de recursos que la reinención de nuestro modelo productivo. En este sentido, la apuesta se ha centrado en un cambio de rumbo con una lógica transversal para la industria y los servicios, con el objetivo de avanzar en una mejor eficiencia y toma de decisiones. Procesos de transformación digital de nuestro tejido productivo que, en definitiva, deben orientarse hacia el incremento de la productividad integrando desarrollos de una IA innovadora en las cadenas de valor de nuestra economía.

Sabemos que la integración de la IA en el tejido industrial y empresarial, al servicio de bienestar y el progreso social y económico, pasa por abordar un trabajo de sinergias y cooperación en el que todos los sectores productivos, empresas de distinto tamaño y administraciones públicas, están llamados a colaborar. Por eso, entendemos la apuesta por la IA en un sentido generacional, como un proyecto de país para la próxima década de cara a alcanzar un crecimiento robusto, sostenible y resiliente basado en potenciar al máximo las nuevas oportunidades de mercado y modelos de negocio.

Este cambio de paradigma a todos los niveles implica ahondar en una visión global sobre cómo debemos incorporar en nuestro modelo económico y social unas tecnologías de IA compatibles con los derechos de las personas, ética y respetuosa con el medio ambiente. En definitiva, una IA capaz de avanzar de manera sostenible con la vida y los principios y valores de nuestras democracias. Una visión que nos lleva a articular los mecanismos para hacer de ella una herramienta igualitaria, transparente y no discriminatoria que, en definitiva, mejore la vida de las personas.

Es importante igualmente tener en cuenta la IA como una tecnología estratégica para la digitalización de la economía y un elemento clave para otra gran transformación de nuestro siglo: la transición ecológica. El aumento de la eficiencia y la productividad son claves para ello. Es necesario fomentar un ecosistema tecnológico que permita impulsar el desarrollo de los sectores de actividad económica y que estos sectores puedan beneficiarse de grandes ventajas competitivas gracias a los avances en IA.

Atendiendo a lo anterior y con esta visión, el gobierno de España aprobó la Estrategia Nacional de Inteligencia Artificial (ENIA), un plan enmarcado en el Plan de Recuperación,

Transformación y Resiliencia, con un presupuesto de 600 millones de euros orientado al impulso de la IA, tanto a través de la innovación y su desarrollo como industria, así como de su integración en las cadenas de valor de la economía. Se trata de una estrategia cuyos efectos se espera que sean estructurales sobre nuestro tejido productivo y para ello se está realizando una inversión que podría llegar a movilizar hasta 3.000 millones de euros en los próximos años. En 2023 más de 500 millones de euros han sido ejecutados o comprometidos como parte del despliegue de esta Estrategia Nacional de Inteligencia Artificial.

Entre los principales proyectos de la ENIA podemos destacar programas centrados en el impulso del talento y la innovación como motores del desarrollo de la IA, como son las Misiones de I+D en IA, cuyo objetivo es abordar grandes desafíos en áreas estratégicas del país a largo plazo aprovechando el potencial transformador de estas tecnologías y sus oportunidades en materia de innovación disruptiva. La finalidad es impulsar la investigación científica y el desarrollo tecnológico centrado en 5 áreas estratégicas para el país: agricultura, salud, medio ambiente, energía y empleo. En esencia, es un plan que busca aprovechar el potencial transformador y de generar impactos y cambios reales de un campo aun por explotar a nivel de innovación como la IA. A través de estas ayudas, ya han financiado proyectos que, con volúmenes de inversión que van desde los 10 a los 20 millones de euros, utilizan la IA e involucran en su desarrollo a organismos de investigación, grandes empresas y PYMES.

Además, se han implementado instrumentos de cofinanciación de proyectos de investigación y desarrollo en Inteligencia Artificial y otras tecnologías digitales y su integración en las cadenas de valor. Una convocatoria con un doble propósito. Por un lado, impulsar la investigación científica, el desarrollo tecnológico y la innovación en IA; y por otro, fomentar el desarrollo tecnológico para la incorporación de la IA en cadenas de valor.

La Estrategia Nacional de Inteligencia Artificial esta asimismo orientada a impulsar a través de su potencial el que es, junto a la digitalización, el segundo pilar transformador al que se ha conducido el despliegue de los Fondos de Recuperación Next Generation UE: la transición ecológica. En este sentido, la Agenda España Digital 2026 plantea un programa de reformas que redundan en la profundización en sinergias que permitan una retroalimentación entre los procesos de transformación digital y energética. Un proyecto de país en el que la Inteligencia Artificial y sus posibilidades van a tener un papel central.

Según el informe de PwC *How AI can enable a Sustainable Future*, el uso de IA para la transición verde tiene efectos positivos sobre el PIB global (+4,4 % a 2030), el empleo neto (38,2 millones a 2030) y la reducción de emisiones de CO<sub>2</sub> (puede eliminar hasta el 4 %). Y se calcula que la IA podría reducir las emisiones mundiales de gases de efecto invernadero en un 4 % hasta 2030. Hablamos de un potencial para la lucha contra la crisis climática que desde el Gobierno de España se ha potenciado con una apuesta transversal por ambos procesos, al tiempo que hemos puesto en marcha medidas para tratar de paliar al máximo los efectos nocivos para el medio ambiente que pueden generar el uso y desarrollo de tecnologías con un gasto energético como las IAs.

En esta dirección se encamina el Programa Nacional de Algoritmos Verdes. Un proyecto que cuenta con una inversión de 257,7 millones de euros de los fondos europeos Next Genera-

tion EU para el desarrollo de una tecnología respetuosa con el medioambiente, guiada por criterios de sostenibilidad y aplicada al desarrollo de acciones contra el cambio climático: verdes por diseño. Esta iniciativa plantea la creación cátedras universitarias asociadas al área temática de algoritmos verdes y que contempla el desarrollo de un área temática específica sobre «IA e Impacto Ambiental» (en el marco de la Red Española de Excelencia en IA). El objetivo es avanzar en el desarrollo de estándares y herramientas para medir el consumo energético de algoritmos, así como en el impulso a la certificación de servicios de hardware y software sostenibles para adoptar un conjunto de buenas prácticas voluntarias, gracias al Sello de Calidad Green Tech que servirá para reconocer aquellos modelos de IA que cumplan con criterios de sostenibilidad.

La Estrategia Nacional de Inteligencia Artificial contempla además importantes medidas para potenciar el desarrollo de una IA en español para impulsar el uso de nuestra lengua en estas nuevas tecnologías como una excelente oportunidad para generar un ecosistema competitivo que entienda el español a nivel global como recurso y materia prima esencial en el desarrollo de una IA con marca España puntera y con proyección a nivel global. Creemos que el fomento de sistemas basados en nuestra lengua común es un factor imprescindible para el conocimiento digital y el desarrollo económico e industrial como la que estamos viviendo.

En este sentido, la ENIA contempla la puesta en marcha del Proyecto estratégico para la recuperación y transformación económica (PERTE) de la Nueva Economía de la Lengua. Esta iniciativa se basa en la combinación de las nuevas tecnologías con el inmenso potencial de la lengua, para impulsar la cadena de valor de la nueva economía digital del conocimiento en español de forma multisectorial. Un proyecto puntero y vanguardista, que pone en valor a las más de 600 millones de personas que forman parte de la comunidad de hispanohablantes, las potencialidades de esta comunidad en el contexto global al garantizar que la digitalización también se desarrolle en español, y el incremento de la competitividad de las empresas españolas y su potencial de expansión global.

Por último, la Estrategia Nacional de Inteligencia Artificial contempla una visión holística y compacta del conjunto del territorio nacional y las distintas regiones que la integran. Una concepción de la innovación en Inteligencia Artificial como elemento cohesionador y un factor orientado al impulso de la vertebración territorial que desde el Gobierno de España se ha potenciado a través de Redes Territoriales de Especialización Tecnológica (RETECH) articuladas a través de iniciativas temáticas basadas en IA. El objetivo de estos proyectos es alinear y articular iniciativas regionales orientadas a la transformación y especialización digital, impulsando la colaboración y la cooperación entre los distintos territorios a través de proyectos de IA conjuntos y complementarios liderados por una o varias regiones.

Atendiendo a lo anterior la conclusión es clara: España cuenta con una visión compartida y colaborativa de las políticas de transformación digital donde las tecnologías de IA tienen un peso decisivo y fundamental para entender un objetivo de país más competitivo, y con un modelo de crecimiento basado en una mejora de la toma de decisiones, de la eficiencia, la sostenibilidad y de un concepto de productividad del siglo XXI.

## Marco de referencia para la Agencia Española de Supervisión de la Inteligencia Artificial

Teniendo en cuenta todo lo anterior, la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022, recoge, en su Disposición adicional centésimo trigésima, la «creación de la Agencia Española de Supervisión de Inteligencia Artificial». Concretamente, se autoriza al Gobierno a impulsar una Ley para la creación de la Agencia Española de Supervisión de Inteligencia Artificial, en adelante la Agencia, configurada como una Agencia Estatal dotada de personalidad jurídica pública, patrimonio propio y autonomía en su gestión, con potestad administrativa. Es por ello, por lo que tuvo que recogerse, posteriormente, en otra norma con rango de Ley la «creación efectiva» de dicha Agencia, como fue el caso de la Disposición adicional Séptima de la Ley 28/2022, de 21 de diciembre, de fomento del ecosistema de las empresas emergentes, también conocida como Ley de Startups. Esta Disposición sirvió para dotar de Presupuesto a la SEDIA de cara a elaborar la documentación y llevar a cabo los trabajos necesarios para la tramitación y creación de esta, además de para poner en liza la importancia de una entidad perteneciente al sector público institucional en un ámbito como la IA.

La creación de la Agencia da cumplimiento a lo previsto en el Plan de Recuperación, Transformación, y Resiliencia, donde se asientan diez políticas palanca que dan desarrollo a una agenda de inversiones y reformas estructurales que se interrelacionan y retroalimentan para lograr cuatro objetivos transversales: avanzar hacia una España más verde; más digital; más cohesionada desde el punto de vista social y territorial; y más igualitaria. La sexta de estas políticas es el Pacto por la Ciencia y la Innovación, en el que se inserta la Estrategia Nacional de Inteligencia Artificial (Componente 16, Reforma 1, del Plan de Recuperación, Transformación, y Resiliencia).

El objetivo de este Componente es situar a España como un país puntero en Inteligencia Artificial, liderando a nivel mundial el uso de la lengua española en la Inteligencia Artificial; promoviendo la creación de empleo cualificado, tanto estimulando el talento español, como atrayendo talento global; incorporando la Inteligencia Artificial como factor para mejorar la productividad de la empresa española y de la Administración Pública; y fomentando sectores estratégicos para nuestro país como el turismo y la hostelería, la agricultura y alimentación, la movilidad, o la energía. Para ello, el Componente se articula a través de la Estrategia Nacional de Inteligencia Artificial, que es uno de los principales planes de la agenda digital del Gobierno de España (España Digital 2026). Algunos de los retos principales abordados por este Componente guardan relación con el uso masivo de la Inteligencia Artificial en las empresas, particularmente en las pymes, con la creación de repositorios de datos de amplio acceso, y con la inversión pública y privada en innovación en materia de Inteligencia Artificial.

La Estrategia Nacional de Inteligencia Artificial vertebra una reforma estratégica de la economía y la sociedad en la era digital, y prevé el establecimiento de un mecanismo de asesoramiento que permita analizar y verificar el desarrollo y la implementación segura y confiable de tecnologías basadas en Inteligencia Artificial, así como de un marco ético y normativo que garantice el desarrollo de una Inteligencia Artificial fiable ajustada a la le-

gislación nacional y europea, y alineada con los valores y principios característicos del Estado social de Derecho en que se conforma España, con especial atención al respeto de la diversidad y la igualdad.

Además, la Agencia contribuye a cuatro de los seis pilares del Reglamento del Mecanismo de Recuperación y Resiliencia: transición digital; crecimiento sostenible, inclusivo, e inteligente; resiliencia sanitaria, económica, institucional, y social; y políticas para las próximas generaciones. El Mecanismo de Recuperación y Resiliencia constituye el núcleo del Fondo de Recuperación acordado por el Consejo Europeo en julio de 2020 para hacer frente a las consecuencias económicas y sociales provocadas por la pandemia de Covid-19, y cuya finalidad es apoyar la inversión y las reformas en los Estados Miembros para lograr una recuperación sostenible y resiliente, al tiempo que se promueven las prioridades ecológicas y digitales de la Unión Europea.

La creación de la Agencia está fundamentada en la obligación que, en virtud de lo dispuesto en la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, se ha establecido para los Estados miembro de seleccionar una «autoridad nacional de supervisión» que se encargue de supervisar la aplicación y ejecución de lo dispuesto en la mencionada Ley de Inteligencia Artificial, así como de coordinar las actividades encomendadas a dichos Estados miembros, actuar como el punto de contacto único para la Comisión, y representar al Estado miembro en cuestión ante el Comité Europeo de Inteligencia Artificial.

La propuesta de Reglamento de Inteligencia Artificial de la Comisión Europea contiene una serie de obligaciones que deberán ser asumidas por la autoridad nacional de supervisión designada. Una regulación cuya filosofía se orienta a la estipulación de un consenso en torno a la limitación de ciertos usos de la IA considerados de «alto riesgo». Es decir, aquellos relativos a aplicaciones que puedan potencialmente atentar contra los derechos fundamentales de las personas, tales como la identificación biométrica, el *scoring social* o aquellas capaces de modificar el comportamiento de las personas de manera subliminal. Para ello, se plantea la creación de una Agencia Estatal, anticipándonos y preparándonos para la asunción de las obligaciones y responsabilidades que imponga el Reglamento, además de otra serie de funciones complementarias que ayudarán a potenciar los sistemas de IA empleados o en desarrollo, así como incorporar una serie de valores que deben incluir los mismos.

La importancia de la creación de una Agencia en el ámbito de la IA queda patente por el creciente uso de esta tecnología en los últimos años, con una previsión futura ascendente aún mayor, así como los importantes beneficios que aportan a las personas cada vez en más ámbitos. Es por ello de donde se desprende su relevancia y necesidad.

Sin perjuicio de lo mencionado anteriormente, tampoco debemos olvidar los riesgos que puede entrañar esta tecnología, dónde entran en liza los derechos recogidos en la Carta de Derechos Digitales, especialmente los relacionados con los ámbitos de la seguridad y salud. Estos riesgos han sido clasificados de menor a mayor intensidad en función de la vulnerabilidad que pueden ocasionar. Una suerte de líneas generales a modo descriptivo y prospec-



tivo que han servido de inspiración para la posterior Declaración Europea sobre los Derechos y Principios Digitales; así como para la elaboración, en colaboración con la Secretaría General Iberoamericana, de la Carta Iberoamericana de Principios y Derechos Digitales.

Estos riesgos derivan de la naturaleza socio-técnica de su aplicación e incluyen, entre otros, amenazas a la seguridad de las personas por posibles resultados imprecisos en su aplicación, o la introducción o amplificación de sesgos que derivan en situaciones de discriminación en el contexto de ciertos casos de uso. Además, ciertos usos de la inteligencia artificial podrían derivar en situaciones de manipulación o polarización de opiniones y acciones de personas. Precisamente en ello estriba la importancia de establecer unos principios en el uso de esta tecnología, así como concienciar, divulgar y formar para poder detectar aquellas conductas que vulneren cualquier derecho. Éste es el objetivo prioritario de la Unión Europea y requiere de los Estados Miembros para la efectividad de las medidas contempladas. En relación a cómo el Reglamento clasifica los sistemas de IA, según su nivel de riesgo se indican algunos ejemplos:

- Un sistema de IA inaceptable sería aquel que manipulara el comportamiento humano de forma subliminal o que explotara a las personas vulnerables, como un juego de azar que indujera a la adicción.
- Un sistema de IA de alto riesgo sería aquel que tuviera un impacto directo sobre la vida o la salud de las personas, como un robot cirujano que operara a un paciente o un sistema de conducción autónoma que controlara un vehículo.
- Un sistema de IA de riesgo limitado sería aquel que generara o modificara contenido de forma automática, como una aplicación de reconocimiento facial que creara un avatar personalizado o un asistente virtual que respondiera a las consultas de los usuarios.
- Un sistema de IA de riesgo mínimo sería aquel que tuviera una finalidad lúdica o recreativa, como un juego de mesa que usara IA para crear desafíos o un filtro de realidad aumentada que cambiara el aspecto de una foto.

Por último, la consecución de los objetivos del Plan Digitalización, incluidos dentro del PRTR, requiere establecer un modelo de gobernanza que permita reintroducir «una fórmula organizativa dotada de un mayor nivel de autonomía y de flexibilidad en la gestión, que cuenta con mecanismos de control de eficacia, y que promueve una cultura de responsabilidad por resultados. Un modelo que cuenta con un enfoque organizativo y funcional y con una filosofía subyacente de gestión dirigida al cumplimiento de objetivos que previamente hayan sido fijados de forma concreta y evaluable».

## Principios de actuación, fines y funciones

A tenor de lo anterior resulta imprescindible constituir este organismo público bajo la naturaleza jurídica de Agencia Estatal<sup>1</sup>, toda vez que dicha forma organizativa permite una mayor agilidad administrativa a la hora de gestionar las funciones señaladas, habilitando,

---

1. En el anexo I se recoge el *Análisis de alternativas que se plantearon antes de crear la Agencia*.

sin merma de los controles de la gestión económico-financiera, presupuestaria y contable comunes a otros organismos públicos, formas flexibles de administración bajo dicha Agencia, dotada de autonomía de gestión y funcional para el mejor cumplimiento de los programas propios de su ámbito competencial.

La Agencia regirá su actuación por varios principios de interés general. Además de por unos principios básicos que serán primordiales para desarrollar sus funciones:

El primero es el principio de no discriminación, que es el origen y fundamento de la propia creación de la Agencia, en cuanto pretende eliminar o luchar contra los sesgos que, inevitablemente, están inmersos en los sistemas de IA desarrollados por los humanos, algunas veces implantados de manera voluntaria y otras de forma inconsciente. Así pues, el derecho a la no discriminación algorítmica se recoge, de manera expresa, en el artículo VIII de la Carta de Derechos Digitales. Precisamente con la Agencia se cumple este mandato de garantizar esta protección en el entorno digital. Si bien, ésta no es la única referencia que entronca directamente con los postulados y misiones de la Agencia, sino que también hay otros derechos de la Carta que merecen ser reseñados, entre los que se encuentran: el artículo I referido a los «derechos y libertades en el entorno digital» donde se garantiza la no discriminación en este ámbito; artículo XII relativo a «Brechas de acceso al entorno digital» que dispone que los Poderes públicos llevarán a cabo políticas activas para evitar las brechas de acceso atendiendo a posibles sesgos discriminatorios, entre otros.

Al anterior se suma el principio de transparencia, basado en la presentación de la información sobre las cuentas de manera precisa y completa, además de las obligaciones de buen gobierno. Nos encontramos también con el principio de autonomía, para alcanzar los objetivos, gestionando sus propios medios. El principio de autonomía, orgánica y funcional, es esencial para el buen hacer de la Agencia que tiene como misión, de cara al futuro más cercano, convertirse en una Agencia que aglutine la mayoría de supervisión de la IA en todos los sectores de actividad, con la finalidad de tener una gestión más eficiente desde la AGE y aprovechar los recursos y *expertise* ganados en este tiempo, cumpliendo así con el principio de eficiencia en la asignación y utilización de los recursos públicos.

El principio de independencia técnica se basa en la alta capacitación y profesionalidad del personal al servicio de la Agencia, adaptado a la ética profesional. Se pretende dotar de diferentes perfiles a la Agencia para poder abarcar varios campos: desde la supervisión más técnica de algoritmos o redes neuronales; a la creatividad en el desarrollo y apoyo en la realización de entornos de prueba o *sandbox* en diferentes ámbitos; también perfiles sociológicos para identificar aquellos principios sociales, éticos o humanistas vulnerados o que requieren de protección; y, perfiles jurídicos para garantizar el funcionamiento de la Agencia acorde al ordenamiento jurídico y supervisar aquellos entornos a los que se exime de la aplicación de la normativa para obtener unos resultados de cómo evoluciona ese concreto sistema de IA en el mercado o en la vida real.

Por último, la Agencia se regirá por el principio de cooperación interinstitucional con los diferentes agentes claves del ecosistema, tanto público como privado. Entendido como la búsqueda de sinergias en la colaboración con otras Administraciones Públicas, agentes e

instituciones, públicas o privadas, nacionales e internacionales para el fomento del conocimiento en todos sus ámbitos.

Son fines generales de la Agencia: la supervisión, el asesoramiento, la concienciación y la formación, como líneas generales, para la adecuada implementación de toda la normativa nacional y europea en torno al adecuado uso y desarrollo de los sistemas de inteligencia artificial, más concretamente, de los algoritmos. Todo ello se complementa con una serie de objetivos que pretenden potenciar y dinamizar la creación, desarrollo y uso de sistemas de IA tanto en el sector público como en el sector privado, pretendiendo que sean, en la medida de lo posible, sistemas esencialmente éticos, sostenibles y/o eficientes energéticamente, para lo que se llevarán a cabo diferentes líneas de actuación.

En este sentido, la Agencia tiene como fines y funciones específicas:

- La promoción de entornos de prueba que permitan una correcta adaptación de sistemas innovadores de inteligencia artificial a los marcos jurídicos en vigor, como el impulso de pilotos de pruebas en entornos seguros y reales; soporte a jueces y tribunales en los sistemas de IA.
- La generación de un marco de confianza en este mismo ámbito. Para ello se creará un marco de certificación voluntaria para entidades privadas y fomentará la dinamización de casos de uso de IA para desarrollar soluciones que permitan transitar hacia modelos productivos y sociales sostenibles, determinando el impacto social en la IA y creando de esta manera un observatorio de tendencias, a través del Observatorio de Impacto Social de los Algoritmos.
- La creación de conocimiento, formación y difusión, especialmente en relación con la inteligencia artificial ética y humanista. Se trata de uno de los fines más importantes puesto que la concienciación y la formación en el desarrollo e implantación de esta nueva tecnología es fundamental para la correcta evolución de la sociedad y la industria española. En efecto, no hay mejor forma de garantizar el progreso de la IA que establecer y defender unos valores y principios entendidos y aceptados por todos.

Se entiende que el conocimiento adquirido desde las instituciones debe trasladarse a las entidades privadas, así como al propio sector público, para garantizar una mayor agilidad en el cumplimiento de todas las futuras obligaciones impuestas a todos los Estados miembro en el futuro en este ámbito. En este sentido, debe destacarse el *sandbox* como un entorno controlado de pruebas del Reglamento europeo IA, que se está desarrollando desde la SEDIA con objeto de operacionalizar los principios del Reglamento aplicables a los sistemas de inteligencia artificial de alto riesgo, así como los de propósito general y los modelos fundacionales, produciendo unas guías de implementación basadas en la experiencia real de hasta 12 proveedores de IA. Estas guías se publicarán abiertamente con el objetivo de facilitar la implementación de este Reglamento, en especial para las PYMEs, consiguiendo una reducción de los costes legales y aportando cierta seguridad jurídica a todos aquellos que desarrollan sistema de IA de alto riesgo. Estas guías también las utilizará la Comisión Europea para desarrollar los actos de ejecución que requiere el Reglamento a partir de la base de experiencia española.

En el futuro, la Agencia tendrá la función de supervisión, concretada y de manera preceptiva, de inspección, comprobación, sanción y demás que le atribuya la normativa europea que le resulte de aplicación y, en especial, en materia de inteligencia artificial, garantizando con ello la extensión del cumplimiento, conforme a los parámetros establecidos, de manera generalizada y normalizada de los sistemas de IA creados, desarrollados o implantados y utilizados por las empresas y diferentes Administraciones, logrando así una cultura ética y humanista de la innovación de la IA. Cada vez son más los actores que desarrollan actividades en las que se emplean sistemas de IA. A través de estas funciones, la Agencia podrá prevenir los principales riesgos del uso de los sistemas de inteligencia artificial como pueden ser los accidentes o los malos usos de dichos sistemas.

Como paso previo al desempeño de la función supervisora de la inteligencia artificial, la AESIA prevé convertirse en un potente think tank, centro de análisis, aportando su visión y su experiencia a los debates y las políticas públicas sobre esta materia.

En relación con este centro de análisis la AESIA podría convertirse en un referente y un líder en el ámbito de la IA, tanto en España como en Europa, y contribuir a la generación de conocimiento y de innovación en esta materia. Asimismo, fomentará un espacio de diálogo y de colaboración entre los diferentes actores y sectores involucrados en la IA, y facilitar la participación y la consulta de la sociedad civil en las decisiones sobre la IA.

La AESIA podría seguir algunos ejemplos de otros *think tanks* sobre IA que existen en el mundo, como el *Centre for Data Ethics and Innovation* del Reino Unido, el *AI Now Institute* de Estados Unidos o el *Future of Humanity Institute* de la Universidad de Oxford.

## Diseño institucional e impactos esperados

Uno de los puntos de partida a la hora de crear la Agencia fue la identificación de la forma jurídica idónea que debería revestir la misma, buscando en todo momento aquella figura que ofreciera ventajas de autonomía y de flexibilidad en la gestión. Para ello, se elaboró un análisis de alternativas jurídicas existentes en la normativa vigente, cuyo detalle queda recogido en el Plan Inicial de Actuación de la Agencia. En este estudio se compararon las figuras de Agencia Estatal, Autoridad Administrativa Independiente y Sociedad Mercantil a través de una serie de indicadores cualitativo como por ejemplo el instrumento de creación, la forma de financiación y cuestiones presupuestarias, el proceso de contratación o el régimen de personal entre otros.

Este estudio, así como distintas consultas con expertos nacionales e internacionales y el servicio jurídico descartaron las alternativas objeto de análisis frente a la figura de la Agencia Estatal respuesta adecuada desde el punto de vista jurídico-institucional por, entre otras cosas, ofrecer ventajas de autonomía y de flexibilidad en la gestión no sólo frente al régimen ordinario aplicable a los órganos administrativos directivos sino frente al de otras tipologías existentes en la normativa vigente.

Estos mecanismos de agilidad, flexibilidad y eficiencia en los procesos de gestión permitirán que la Agencia pueda ofrecer una respuesta rápida y efectiva en el diseño, ejecución y

evaluación del alto volumen de proyectos estratégicos de IA, así como adaptarse a los ajustados plazos de ejecución delimitados por la Unión Europea a la hora de gestionar los recursos procedentes del PRTR

Por otro lado, la elección de esta figura jurídica permitirá que, al estar adscrita a la Secretaría de Estado de Digitalización e Inteligencia Artificial, se puedan reutilizar perfiles altamente especializados en IA. Asimismo, también se elaborará una nueva Relación de Puestos de Trabajo con la identificación de perfiles de alta especialización que desarrollen de forma óptima y eficiente las actuaciones de la Agencia, los cuales contarán con conocimientos vanguardistas en esta materia lo que permitirá adaptarse a las modificaciones que, en este ámbito se producen con mucha rapidez. Otros de los beneficios de este modelo serán la convergencia tecnológica; la prestación de servicios eficiente; la sostenibilidad de los servicios digitales o el modelo tecnológico común.

Además, a través del despliegue de sus actuaciones, la AESIA potenciará la innovación teniendo efectos positivos a medio largo plazo de notoria importancia. En este sentido, las actividades públicas deben procurar promover la actividad de I+D+i, así como la incorporación de nuevas tecnologías al proceso productivo. La innovación debe dirigirse a sectores que vayan en línea con nuestros principios y sistemas. En este sentido, se velará por prevenir y evitar que los desarrollos de IA sirvan para actividades ilícitas tales como el blanqueo de dinero, financiación del terrorismo o evasión de impuestos.

El impacto que tendrá la AESIA por razón de género es positivo, ya que se regirá en el ejercicio de sus funciones específicas por el principio de igualdad de género, promoviendo la perspectiva de género y una composición equilibrada de mujeres y hombres en sus órganos, consejos y actividades de acuerdo con lo dispuesto en la Ley Orgánica 3/2007, de 22 de marzo, para la igualdad efectiva de mujeres y hombres.

Se considera que el impacto por razón de cambio climático es positivo, ya que la AESIA impulsará medidas relacionadas con el apoyo al desarrollo y uso de sistemas de Inteligencia Artificial desde una perspectiva ambiental.

En otro sentido, a medio largo plazo la AESIA podrá orientar actividades vinculadas a la coordinación de otros servicios, ámbitos y actividades que puedan derivarse de la diferente normativa europea, incluida la supervisión en diferentes campos, más allá del Reglamento de Inteligencia Artificial.

## Reflexiones finales

La Agencia Española de Supervisión de Inteligencia Artificial es la primera institución de estas características que se crea en la Unión Europea. La Agencia Española es pionera y se anticipa a todas aquellas obligaciones que, desde la Unión Europea, se impongan en el despliegue, implementación y desarrollo de los sistemas de IA. La finalidad de la Agencia es dar certidumbre a estos sistemas, que se guíen por una serie de principios uniformes, así como facilitar el cumplimiento de la futura normativa en este ámbito a las empresas y entidades que la emplean.

En lo relativo a las competencias, la creación de esta Agencia Estatal tiene como objetivo la supervisión de los sistemas de inteligencia artificial empleados tanto por el Sector Público como por el Sector Privado. Esta Agencia asumirá, entre otras, las obligaciones y responsabilidades que imponga el futuro Reglamento de IA, además de absorber los resultados que se obtengan del entorno de cumplimiento de requisitos de sistemas de IA y tiene como objetivo ser la base del citado Reglamento europeo, pudiendo considerar a la AESIA como una figura de apoyo a la Unión Europea para dictar legislación. En este sentido, al impulsarse este piloto a nivel nacional y ser la AESIA quien absorberá los resultados de la iniciativa existe la posibilidad que sea la Agencia la encargada de realizar tanto posteriores entornos controlados de prueba, como de modificar los textos normativos en relación a los cambios que puedan existir en la supervisión de la IA o en las redes neuronales.

En efecto, los organismos de estandarización ya están elaborando las normas técnicas para la aplicación de dicho Reglamento que, a su vez, se está redactando en el seno de la Comisión Europea, Parlamento Europeo y Consejo Europeo, con el objeto de aprobar las normas armonizadas en materia de IA. Para ello, la Agencia tiene una labor importante en materia de coordinación de los trabajos y resultados del mencionado entorno controlado, creación de entornos controlados de prueba análogos o complementarios, así como, garantizar la correcta aplicación de la normativa europea en materia de Inteligencia Artificial.

## Referencias

Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022.

Ley 31/2022, de 23 de diciembre, de Presupuestos Generales del Estado para el año 2023.

Carta de Derechos Digitales, Plan de Recuperación, Transformación y Resiliencia del Gobierno de España.

Plan de Recuperación, Transformación y Resiliencia del Gobierno de España.

Estrategia Nacional de Inteligencia Artificial, ENIA. (2021.) Consultado a través de <https://portal.mineco.gob.es/es-es/digitalizacionIA/Paginas/ENIA.aspx>, en mayo de 2021

Ley 28/2022, de 21 de diciembre, de fomento del ecosistema de las empresas emergentes, Disposición adicional séptima. «Creación de la Agencia Española de Supervisión de Inteligencia Artificial».

Documentación de la Audiencia pública sobre el Real Decreto que establece un entorno controlado de pruebas para el ensayo del cumplimiento de la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial. Consultado en [https://portal.mineco.gob.es/es-es/ministerio/participacionpublica/audienciapublica/Paginas/audiencia\\_entorno\\_controlado\\_pruebas\\_ensayo\\_cumplimiento\\_normas\\_armonizadas\\_IA.aspx](https://portal.mineco.gob.es/es-es/ministerio/participacionpublica/audienciapublica/Paginas/audiencia_entorno_controlado_pruebas_ensayo_cumplimiento_normas_armonizadas_IA.aspx), en agosto de 2023

Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial.

Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión (COM/2021/206 final).