

PANORÁMICA INTERNACIONAL DE LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL: REGLAMENTO EUROPEO DE INTELIGENCIA ARTIFICIAL (AI ACT)

Juan Pedro Díaz Senés*

Investigador predoctoral de Derecho civil

Universidad Rey Juan Carlos

Resumen: Durante el año 2023, los países con mayor desarrollo tecnológico han realizado una importante actividad regulatoria en materia de Inteligencia Artificial, acentuándose la voluntad de los Estados para alcanzar acuerdos que delimiten el riesgo de estos sistemas. Se analizarán las propuestas de la UE, en concreto el Reglamento de Inteligencia Artificial, y del G-7, conocido como Proceso de Hiroshima, así como la celebración de la primera Cumbre de Seguridad de la IA, en Reino Unido.

Palabras clave: Inteligencia artificial, riesgo, RIA, Proceso de Hiroshima, Cumbre de Seguridad.

Title: International overview of the regulation of Artificial Intelligence. European Artificial Intelligence Act

Abstract: In the year 2023, the main countries that lead technological development have carried out important regulatory activity in the field of Artificial Intelligence, accentuating the desire of States to reach agreements that delimit the risk of these systems. The EU proposals, specifically the Artificial Intelligence Regulation, and the G-7, known as the Hiroshima Process, will be analyzed, in addition to the holding of the first AI Security Summit, in United Kingdom.

Keywords: Artificial Intelligence, risk, Hiroshima Process, AI Act, Security Summit.

SUMARIO: 1. Cuestiones generales. 2. Acuerdos europeos para la aprobación del Reglamento de Inteligencia Artificial. 3. Cumbre del G-7. El Proceso de Hiroshima. 4. Cumbre de seguridad de la IA en Reino Unido. 5. Conclusiones.

1. CUESTIONES GENERALES

En escasos cinco años la Inteligencia Artificial (IA) ha dejado de ser un concepto utilizado de forma exclusiva en el campo científico para ocupar prácticamente todos los ámbitos de la sociedad. Es común encontrar numerosas noticias en las que se alude a los grandes beneficios que trae consigo esta tecnología. Como son aquellas en las que se describen herramientas que permiten redactar o mejorar un escrito, la posibilidad de crear una imagen o vídeo a partir de meras ideas o los avances en vehículos inteligentes, con la expectativa de ser en un futuro plenamente autónomos².

En contraposición, son aún más numerosos los titulares en medios de comunicación y sitios webs de divulgación científica que advierten con ímpetu de los grandes riesgos a los que se enfrenta la sociedad. Por ejemplo, este titular publicado en el mes de mayo de 2023 por el diario El País: “Los principales creadores de la IA alertan sobre el “peligro de extinción” que supone esta tecnología para la humanidad”³. Es de suma importancia en este contexto definir y acotar qué es la Inteligencia Artificial, qué riesgos se derivan de su uso y cuáles son las previsiones sobre su futuro desarrollo. Desde la perspectiva del Derecho, las respuestas a estas cuestiones permitirán adaptar el ordenamiento jurídico a las necesidades que trae consigo esta tecnología emergente.

Las instituciones de la Unión Europea conscientes de ello están ultimando su regulación para la aplicación y uso de los sistemas de IA, la cual, a falta de su aprobación, en este año 2024, se presenta como la primera norma sobre esta materia en el mundo⁴. Así, más allá de medidas concretas impulsadas por países como EE.UU.⁵ o China⁶, son de especial interés los acuerdos de instituciones y foros

² La conducción autónoma se clasifica conforme a una escala de 0 a 5, correspondiendo los valores 4 y 5 a una conducción autónoma. Actualmente, son numerosos los países que permiten la conducción autónoma en espacios públicos. En el caso de España, se está trabajando en un Real Decreto que permita la circulación de esta modalidad de vehículos para este año, 2024.

³ El País, “Los principales creadores de la IA alertan sobre el “peligro de extinción” que supone esta tecnología para la humanidad”. Disponible en: <https://elpais.com/tecnologia/2023-05-30/los-principales-creadores-de-la-ia-alertan-sobre-el-peligro-de-extincion-que-supone-esta-tecnologia-para-la-humanidad.html#>

⁴ Consejo europeo, “Reglamento de Inteligencia Artificial: el Consejo y el Parlamento alcanzan un acuerdo sobre las primeras normas del mundo en materia de inteligencia artificial”. Disponible en: <https://www.consilium.europa.eu/es/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>

⁵ El 30 de octubre se anunció una orden ejecutiva por parte del gobierno estadounidense en materia de inteligencia artificial. Esta regulación impone medidas de transparencia a las empresas tecnológicas que operen en EE. UU. Disponible en: <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>

⁶ En el mes de agosto la República Popular de China aprobó dos nuevas normas en materia de inteligencia artificial, con un enfoque basado en el riesgo de forma similar a las propuestas de la UE y el G-7. Para un

intergubernamentales, dada su capacidad para influir en el resto de los ordenamientos extranjeros.

Especialmente sobre esta materia la incidencia es aún mayor que en otros ámbitos, debido a que el desarrollo y los avances en los sistemas de IA no se detendrán en las fronteras de un Estado u otro, sino que previsiblemente se implantarán en cualquier mercado del mundo. Piénsese en la cantidad de sistemas de IA que son únicamente un software y no requieren de ningún dispositivo físico, por ejemplo, como ya se ha visto en los sistemas de IA generativas, que crean una imagen o modifican el contenido de un vídeo. Estos modelos apenas presentan dificultades para su comercialización en cualquier lugar, por tanto, para minimizar los riesgos que puedan traer consigo estos sistemas es más que deseable una regulación que sea de aplicación generalizada.

En este sentido, en el año 2023 se celebraron dos importantes acuerdos y dos cumbres internacionales en torno a la regulación de los sistemas de IA. Fueron a) los acuerdos tripartitos de la UE, en diciembre, para la aprobación del Reglamento en materia de inteligencia artificial (RIA), b) la Cumbre del G-7 en Japón durante el mes de mayo, que propició los Acuerdos de Hiroshima y c) la Cumbre de seguridad mundial en torno a la Inteligencia Artificial, "AI Safety Summit", en el mes de noviembre en Reino Unido.

2. ACUERDOS EUROPEOS PARA LA APROBACIÓN DEL REGLAMENTO DE INTELIGENCIA ARTIFICIAL

En abril de 2021, se presenta 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, y que abordará los principales retos y oportunidades en este ámbito (COM 2021, 206 final) (Propuesta del RIA). Esta propuesta obedece a la voluntad de constituir un adecuado marco regulatorio, conforme a los estándares de seguridad, confianza y ética. Como resultado de los acuerdos celebrados a lo largo del año 2023⁷, se ha anunciado su aprobación en este año 2024⁸.

En este tiempo, además de esta propuesta, se ha trabajado en la elaboración de nuevos modelos armonizados sobre responsabilidad civil por los daños ocasionados

análisis sintetizado de ambas normas, véase a B. Fernández, C. "Entra en vigor la Ley china sobre algoritmos de recomendación en Internet" *Derecho Digital e Innovación. Digital Law and Innovation Review*, núm. 13, 2022.

⁷ En 14 de junio, el Parlamento introdujo una serie de modificaciones sobre la propuesta y posteriormente los días 8 y 8 de diciembre se celebraron los conocidos como diálogos tripartitos por parte del Parlamento, Consejo y Comisión, que concluyó con el acuerdo final del texto.

⁸ Consejo europeo, "Reglamento de Inteligencia Artificial: el Consejo y el Parlamento..." op. cit.

por la IA⁹ y la adaptación de las normativas en otros ámbitos donde inciden estos sistemas. Entre todas ellas, se prevé que, en 2024, sea aprobada una nueva Directiva sobre responsabilidad por los daños causados por productos defectuosos¹⁰, adaptando el contenido de la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985 a los elementos de la economía digital.

El RIA está previsto que se articule como la norma de referencia, con un carácter general y estructura horizontal, frente al resto de normas sectoriales y disposiciones internas de los Estados Miembros. Al igual que ya sucedió con el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (RGPD), respecto al derecho fundamental a la protección de los datos de carácter personal.

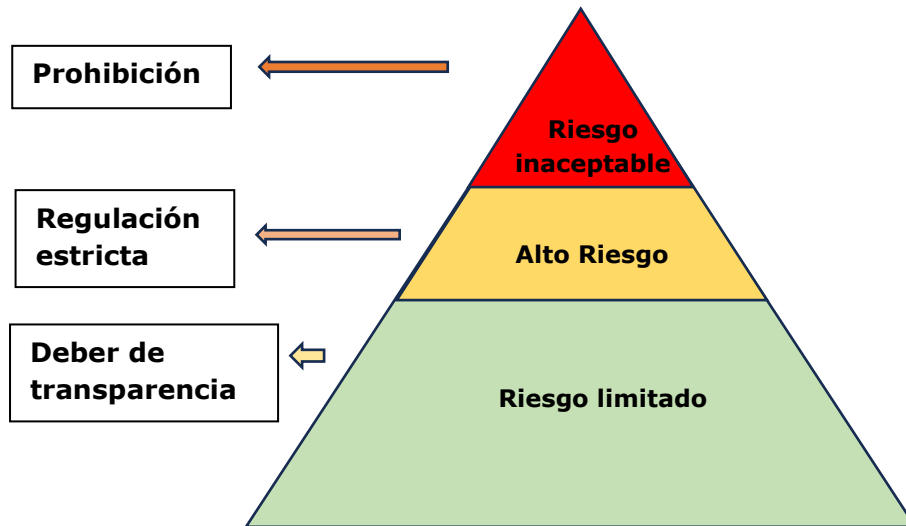
Los objetivos que se plantean a través de este Reglamento son la constitución de un marco jurídico sólido, al mismo tiempo que se incentivará la innovación tecnológica. A priori, son estos intereses contrapuestos pues una protección sumamente garantista irá en detrimento del desarrollo tecnológico. En este sentido, el RIA pivota en torno a la esfera de la protección de los derechos fundamentales europeos, siendo clasificados los sistemas de IA conforme al riesgo que pueden generar, tanto a un individuo como a todo un colectivo o a una infraestructura.

Esta clasificación de los sistemas de IA responde a un enfoque de gestión del riesgo. Los sistemas de riesgo bajo o mínimo, es decir aquellos con escasa capacidad para inferir un daño, se les impondrán reducidas medidas de control. Por el contrario, los sistemas de alto riesgo o sistemas prohibidos, que entrañan riesgos elevados o inasumibles, se les dedica el grueso de la regulación¹¹.

⁹ Han sido publicados dos propuestas de regulación en materia de responsabilidad extracontractual por daños causados por IA, con iniciativas muy distintas: la Propuesta de Reglamento en materia de responsabilidad civil por el uso de inteligencia artificial, del Parlamento Europeo, de 20 de octubre de 2020, <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020IP0276&from=EL> y la Propuesta de Directiva del parlamento europeo y del consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA) de 28 de septiembre de 2022. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52022PC0496>.

¹⁰ Consejo europeo: "Un acuerdo entre el Consejo y el Parlamento Europeo adapta las normas de la UE sobre responsabilidad a la era digital y a la economía circular". Disponible en <https://www.consilium.europa.eu/es/press/press-releases/2023/12/14/agreement-between-the-council-and-the-european-parliament-makes-eu-liability-rules-fit-for-the-digital-age-and-circular-economy/>

¹¹ Sobre la gestión de riesgos y los daños ocasionados en el contexto tecnológico, *vid* Jimeno Muñoz, J. *Derecho de daños tecnológicos, ciberseguridad e insurtech*. 1 ed. Dykinson. Madrid, 2019.



Fuente: Briefing AI Act. Parlamento Europea

Esta sistematización sobre los sistemas de IA conforme al estándar de protección de los derechos fundamentales es una medida garantista, pues obedece a la capacidad de los sistemas de causar un daño adecuándose a unos principios bien reconocidos en el marco normativo europeo (la Carta de los Derechos Fundamentales de la UE). No obstante, el aspecto que más debate ha generado a lo largo del proceso de tramitación del RIA ha sido delimitar que que sistemas quedan en uno u otro nivel. Es decir, acordar que sistemas deben ser identificados como riesgos prohibidos, no pudiendo ser comercializados en el mercado, y cuales aun estando permitidos requieren de fuertes controles. La postura sostenida por el Parlamento, en ocasiones, ha sido opuesta a la mantenida por el Consejo, especialmente en las restricciones que sobrevuelan al sistema de vigilancia e identificación biométrica.

En orden cronológico, tras la publicación de la Propuesta del RIA en 2021 se alcanzó el acuerdo para el inicio del mandato de negociación entre las Comisión, el Consejo y el Parlamento, en diciembre de 2022 (diálogos tripartitos). Posteriormente, en junio de 2023 el Parlamento aprueba una serie de sustanciales modificaciones sobre el contenido de la Propuesta, dotándolo de mayor protección para los ciudadanos de la UE, y finalmente en diciembre de 2023 se alcanza, tras nuevos cambios en el texto, el esperado acuerdo entre las tres instituciones, con lo que el texto se consolida de forma definitiva, a falta de su publicación en el año 2024.

Las modificaciones que se han introducido en el RIA atienden al concepto de inteligencia artificial y a otros conceptos asociados al mismo, esenciales para delimitar su ámbito de aplicación. Asimismo, se han añadido nuevas disposiciones para regular la aplicación de los modelos fundacionales y reglas de transparencia sobre los modelos generativos.

La definición por la que se ha optado finalmente para el término inteligencia artificial obedece al propuesto por la Organización para la Cooperación y el Desarrollo Económicos (OCDE), que manifiesta la voluntad de la Unión Europea de colaborar con el resto de agentes internacionales implicados. Así, se define la inteligencia artificial como:

“Un sistema basado en máquinas diseñado para funcionar con diversos niveles de autonomía y capaz, para objetivos explícitos o implícitos, de generar información de salida –como predicciones, recomendaciones o decisiones– que influya en entornos reales o virtuales”.

También se han introducido una serie de principios estructurales común a cualquier modelo de inteligencia artificial¹² y se han añadido nuevos términos como “riesgo significativo”, “modelo fundacional” (el ejemplo más conocido es el modelo ChatGPT), “ultrafalsificación”, o “espacio controlado de pruebas”. Se caracterizan todas ellas por utilizar en su vocabulario términos amplios y conceptos indeterminados. Si bien, se ha de valorar positivamente cierta abstracción en los conceptos, pues la utilización de términos concretos genera el riesgo de su inaplicación temprana¹³, dado los rápidos avances que se están produciendo en esta tecnología y la diversidad de usos.

De vuelta con la clasificación de los sistemas de IA, los acuerdos alcanzados en el mes de diciembre entre el Parlamento y el Consejo sitúan finalmente como uso inaceptable, y, por tanto, prohibidos, aquellos sistemas destinados a:

- La manipulación cognitiva conductual
- El rastreo indiscriminado de imágenes faciales, obtenidas de sitios webs y circuitos cerrados de televisión
- El reconocimiento de emociones en espacios de trabajo e instituciones de enseñanza
- La puntuación ciudadana y la categorización biométrica para obtener datos como orientación sexual, ideología, creencias religiosas
- La vigilancia e identificación biométrica de personas

Como se indicaba con anterioridad, las posturas a favor y en contra de restringir el uso de sistemas de identificación biométrica han sido objeto de importantes debates. El Parlamento se ha mostrado partidario de prohibir los sistemas de identificación biométrica remota cuando suceda en tiempo real y en espacios de acceso público y,

¹² Son: intervención y vigilancia humanas; solidez y seguridad técnicas; privacidad y gobernanza de datos; transparencia; diversidad, no discriminación y equidad; y bienestar social y medioambiental. Todos ellos representan una concreción de los principios generales presentes en la Carta de los Derechos Fundamentales de la Unión Europea (2000/C 364/01).

¹³ En este mismo sentido, Barrio Andrés M. “Novedades en la tramitación del próximo Reglamento europeo de inteligencia artificial”. Disponible en: <https://www.realinstitutoelcano.org/analisis/novedades-en-la-tramitacion-del-proximo-reglamento-europeo-de-inteligencia-artificial/>.

además, los sistemas de identificación biométrica remota en diferido, salvo cuando sea utilizado para la persecución de delitos graves previa autorización judicial.

La posición mantenida por el Parlamento ha sido apoyada por otras instituciones como la Organización Europea de Consumidores o la agrupación *European Digital Rights*. Se pretende así impedir un control sistematizado y masivo sobre los ciudadanos europeos, como sucede, por ejemplo, en China¹⁴. Sin embargo, las negociaciones mantenidas en el mes de diciembre matizaron estas restricciones, permitiendo que los sistemas de identificación biométrica remota, tanto en tiempo real como en diferido, puedan ser usados por parte de determinadas autoridades y en contextos estrictamente necesarios. Por tanto, se ha matizado tales prohibiciones, en aras de salvaguardar la seguridad nacional e impedir la consecución de delitos graves.

Respecto a otra importante novedad que introducen los acuerdos de 2023, se ha añadido una regulación pormenorizada sobre los modelos fundacionales. En la propuesta inicial del RIA se carecía de normas concretas sobre estos sistemas, al no intuir la UE la gran evolución y riesgos que su implantación introdujo en la sociedad. Estos sistemas se caracterizan por su capacidad para generar texto, sonido, códigos informáticos, videos e imágenes. El modelo más conocido actualmente es ChatGPT o Meta's Llama.

Las medidas de control previstas para los modelos fundacionales consisten en obligaciones concretas en materia de transparencia previas a su introducción en el mercado, tales como impedir que puedan genere contenido ilegal, o que mediante su actividad vulneren derechos de autor e informar aquellos contenidos que han sido generados por ellos. La capacidad de esta tecnología es actualmente muy alta y se prevé que progresivamente mejore, dificultando en gran medida identificar que se ha generado mediante IA y que no. Para estos sistemas fundacionales que tienen capacidad para ocasionar riesgos sistémicos está previsto un control más estricto.

Sin intención de ser exhaustivo en la enunciación de los riesgos de los modelos fundacionales, piénsese en un audio, foto o video que sitúa a una persona donde nunca ha estado o realizando un hecho que no existió, si no es posible identificar la falsedad de tales sonidos o imágenes el daño para la víctima no podrá ser reparado. En otros ámbitos también existen riesgos reseñables, como podría ser la elaboración de un trabajo académico (trabajo fin de grado, tesis doctoral o un artículo científico), que tiene como "autor" un sistema fundacional, en este escenario si no existen

¹⁴ En noviembre de 2023, China ha presentado un nuevo proyecto para la regulación del *Social Credit System*, modelo que tiene por objeto recompensar las acciones que generan confianza en la sociedad y penalizan aquellas que no, mediante una supervisión masiva de la población. Entre los principales detractores, Brussee V., *Social credit: the warring states of China's emerging data empire* ed. electr. Singapore : Palgrave Macmillan, 2023.

medios para verificar la autenticidad del trabajo habrá que replantearse qué objeto tiene su elaboración.

Aparte de las restricciones que contiene el RIA, se está construyendo en paralelo una nueva estructura de gobernanza, en la que distintas entidades trabajarán de forma coordinada para garantizar que las constantes evoluciones en esta tecnología no repercutan negativamente en la protección de los derechos de los ciudadanos europeos. En concreto, se creará en el año 2024 una oficina de IA, adscrita a la Comisión con expertos independientes para su asesoramiento, un Comité de IA, compuesto por representantes de los Estados miembros y un foro consultivo con la finalidad de aglutinar los conocimientos de todos los agentes involucrados: sector industrial, sociedad civil y académicos.

Por último, se han aumentado las sanciones en caso de incumplimiento, respecto a la propuesta de RIA de 2021:

- Implantación de sistemas prohibidos: 35 millones de euros o el 7% del volumen de negocio del año anterior de la empresa infractora, si este fuera superior
- Incumplimiento de las obligaciones contenidas en la norma: 15 millones de euros o el 3% del volumen de negocio, si este fuera superior
- Presentación de información inexacta: 7,5 millones de euros o el 1,5 % del volumen de negocio, si este fuera superior

En todo caso, las sanciones podrán ser de una cuantía inferior, en proporción al tamaño de la empresa y a la gravedad de la disposición infringida. La autoridad que sea designada para la supervisión de los sistemas, se prevé que disponga de un procedimiento específico para que puedan presentar reclamaciones las personas físicas y jurídicas, al igual que ya sucede en el RGPD.

3. CUMBRE DEL G-7. EL PROCESO DE HIROSHIMA

Más allá de los notables esfuerzos de la UE por liderar el proceso de regulación de la IA no se ha de obviar la actuación del resto de países y agentes involucrados, como son los acuerdos del G-7 sobre esta materia.

Los días 19 a 21 de mayo de 2023, se reunieron los integrantes del G-7 en Hiroshima. En esta cumbre se manifestó el compromiso por avanzar en los debates ya iniciados en el mes de abril¹⁵ sobre la gobernanza de la IA, en consonancia con los valores democráticos manifestados en la Carta de las Naciones Unidas. Como resultado de estos debates, en el mes de septiembre de este mismo año se publicó el Informe de

¹⁵ Comunicado de líderes G-7 Cumbre de Hiroshima: <https://www.consilium.europa.eu/media/64497/g7-2023-hiroshima-leaders-communic%C3%A9.pdf>.

trabajo sobre la Inteligencia Artificial Generativa¹⁶ y en octubre se anunció un nuevo compromiso por parte de sus miembros, que se conoce como el Proceso de IA de Hiroshima.

El Proceso de IA de Hiroshima plantea una vía normativa distinta a la propuesta por la Unión Europea. Tapia Hermida¹⁷ describe este proceso como un “sendero medio” o “tercera vía”, en la que se asume la necesidad de construir mecanismos para garantizar que esta tecnología sea desarrollada en términos de fiabilidad, pero sin asumir las directrices de la Unión Europea, que cabrían ser descritas como imperativas o propias de un modelo de *strict regulations*¹⁸. En él se reconoce, por una parte, el cumplimiento de los Principios Rectores Internacionales en materia de IA elaborado por la OCDE¹⁹ y, por otra, la publicación Código de conducta voluntario para los desarrolladores de Inteligencia Artificial²⁰, en línea con estos principios.

Los Principios Rectores Internacionales en materia de IA elaborado por la OCDE tienen como destinatarios todos aquellos actores que interactúan con un sistema de estas características: diseñadores, desarrolladoras, implementadores, proveedores, importadores y usuarios. Asimismo, no se comportan como una lista cerrada, pudiendo ser extensiva a otros sujetos o complementarse a través de otros principios externos. Tapias Hermida identifica cada uno de estos principios con las distintas fases del control y gestión del riesgo, derivado del uso de esta tecnología: “gestión de los riesgos”, “prevención de los incidentes” y “solución de incidentes”²¹. El contenido de los principios, preservando el orden en el que son enunciados en el texto oficial, son:

1. *Tomar medidas apropiadas para identificar, evaluar y mitigar los riesgos antes y durante el despliegue;*

¹⁶ OECD, “G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI”. Disponible en: <https://doi.org/10.1787/bf3c0c60-en>

¹⁷Tapia Hermida, A. “Inteligencia Artificial Responsable. Dos documentos relevantes para su regulación: El Informe Hiroshima sobre la Inteligencia Artificial Generativa y el Proyecto internacional de principios rectores para las organizaciones que desarrollan sistemas avanzados de Inteligencia Artificial”. En: <https://ajtapia.com/2023/11/inteligencia-artificial-responsable-dos-documentos-relevantes-para-su-regulacion-el-informe-hiroshima-sobre-la-inteligencia-artificial-generativa-y-el-proyecto-internacional-de-principios-rectores-p/>.

¹⁸ Daisuke T. “New global norms on responsible AI beyond the EU: the G7 Hiroshima Process international guiding principles for developing advanced AI systems and its application in Japan”. En: https://www.ibanet.org/G7-Hiroshima-Process-international-guiding-principles-for-developing-advanced-AI-systems-application-in-Japan#_ftn5.

¹⁹ Principios Rectores Internacionales en materia de IA elaborado por la OCDE: <https://digital-strategy.ec.europa.eu/es/library/hiroshima-process-international-guiding-principles-advanced-ai-system>

²⁰ Código de conducta voluntario para los desarrolladores de Inteligencia Artificial: <https://ec.europa.eu/newsroom/dae/redirect/document/99641>.

²¹ En este sentido, Tapias Hermida en la primera fase (la gestión de los riesgos) incluye los principios primero, tercero, quinto, noveno y décimo; en la segunda fase (prevención de los incidentes), señala los principios segundo, sexto, séptimo, octavo y undécimo; y por último en la tercera fase (solución de incidentes) indica el principio número cuarto. Tapias Hermida A. “Inteligencia Artificial...” op. cit.

2. *identificar y mitigar vulnerabilidades, incidentes y patrones de uso indebido después del despliegue;*
3. *informar públicamente las capacidades, limitaciones y dominios de uso apropiado e inapropiado para la transparencia y la rendición de cuentas;*
4. *trabajar para compartir información y notificar incidentes de forma responsable;*
5. *desarrollar, implementar y divulgar políticas de gestión de riesgos y gobernanza de la IA, basadas en un enfoque basado en el riesgo;*
6. *invertir e implementar controles de seguridad sólidos;*
7. *desarrollar e implementar mecanismos confiables de autenticación y procedencia de contenido;*
8. *priorizar la investigación para mitigar los riesgos y la inversión en medidas de mitigación efectivas;*
9. *priorizar el abordaje de los desafíos mundiales, como la crisis climática, la salud y la educación globales;*
10. *avanzar en el desarrollo y adopción de normas técnicas internacionales;*
11. *implementar medidas de entrada de datos y protecciones para datos personales y propiedad intelectual.*

Esta clasificación, como se recoge en su preámbulo, presenta una estrecha relación con otros estándares internacionales previos, en concreto los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos²² o las Directrices de la OCDE para Empresas Multinacionales²³. A su vez, los Principios Rectores Internacionales en materia de IA se emplean como fundamento para la formulación del Código de conducta voluntario de Inteligencia Artificial del G-7.

El Código de Conducta de la Inteligencia Artificial es una disposición que sugiere e invita a cualquier Estado, así como a quienes intervienen en el ciclo de vida de los sistemas, a asumir ciertos comportamientos y medidas con los que "promover una IA segura y digna de confianza en todo el mundo y proporcionar una guía voluntaria para las acciones de las organizaciones que desarrollan los sistemas de IA más avanzados, incluidos los modelos de cimentación y los sistemas generativos de IA más avanzados"²⁴ Por tanto, es una regulación de carácter *soft law*, en la que no hay un mandato de obligado comportamiento ni normas de carácter imperativas, como sí sucede en el RIA de la UE. Tanto es así, que en este Código se indica que "que los distintos ordenamientos jurídicos podrán adoptar su propia orientación para aplicar estas medidas de distinto modo"²⁵.

²² ONU: Oficina del Alto Comisionado de Derechos Humanos (ACNUDH), ONU: "Principios Rectores sobre las empresas y los derechos humanos: puesta en práctica del marco de las Naciones Unidas para 'proteger, respetar y remediar'". En: <https://www.refworld.org/es/docid/5d7147f14.html>.

²³ OCDE (2013), Líneas Directrices de la OCDE para Empresas Multinacionales, OECD Publishing. En: <http://dx.doi.org/10.1787/9789264202436-es>.

²⁴ "Código de conducta voluntario..." p. 1.

²⁵ "Código de conducta voluntario..." p. 8.

Con relación a su contenido se sustenta, como ya se incidió, en los Principios Rectores Internacionales en materia de IA, representando sobre cada uno de estos principios posibles manifestaciones prácticas. En concreto, consiste en una "lista no exhaustiva de acciones que se discute y elabora como un documento vivo"²⁶. Estas acciones se comportan como un modelo de análisis coste(riesgo)/beneficio²⁷, asumiendo como objetivo construir una IA segura y digna de confianza, en sintonía plena con la finalidad de la Propuesta de Ley de IA de la UE.

Por último, atendiendo a su forma y fondo el Código de Conducta presentado por el G-7 es compatible con la propuesta de RIA de la UE, siendo posible la aplicación de ambas disposiciones de modo complementario. Von der Leyen, presidenta de la Comisión Europea, ha manifestado con entusiasmo: "me complace dar la bienvenida a los Principios Rectores internacionales del G-7 y al Código de Conducta voluntario, que reflejan los valores de la UE para promover una IA fiable. Pido a los desarrolladores de IA que firmen y apliquen este código de conducta lo antes posible"²⁸.

4. CUMBRE DE SEGURIDAD DE LA IA EN REINO UNIDO

Con una participación notablemente superior que los acuerdos del G-7 y de la UE, los días 1 y 2 de noviembre de 2023 se celebró la cumbre de seguridad de la IA en Reino Unido. En él, fueron citados numerosos miembros de la industria tecnológica, expertos y representantes de distintos Estados. Entre las autoridades, destaca la presencia de la Unión Europea, que ocupó un espacio relevante en este foro.

Además de las propuestas regulatorias que traían consigo los Estados²⁹, se desprende de esta cumbre la voluntad de cumplir un objetivo común, motivado por una misma preocupación: delimitar los riesgos asociados con la IA, especialmente cuando repercutan en sectores vulnerables, como la ciberseguridad o la defensa fronteriza. La Declaración de los países que asistieron a la Cumbre de Seguridad de La IA (la Declaración)³⁰, incidió en los riesgos singulares que se presenta en la "frontera de la

²⁶ "Código de conducta voluntario..." p. 1.

²⁷ Hermida Tapia explica con detalla como los distintos principios inciden en las tres fases del tradicional análisis de gestión de riesgo. Tapia Hermida, A. "El Código de Conducta de la Inteligencia Artificial (CCIA) del G-7". En: <https://ajtapia.com/2023/11/el-codigo-de-conducta-de-la-inteligencia-artificial-ccia-del-g-7/>.

²⁸ Comisión europea. "La Comisión se congratula del acuerdo de los dirigentes del G-7 sobre principios rectores y un código de conducta en materia de inteligencia artificial". En: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5379

²⁹ Véase a nota 5.

³⁰ The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023. En: <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

IA". Este concepto, "frontera de la IA" alude a aquellos modelos capaces de realizar una amplia variedad de tareas, así como IA específicas que son capacidades de causar un daño, que iguale o supere las capacidades actuales en los modelos más avanzados de hoy.

En definitiva, la "frontera de la IA" se identifica fácilmente como sinónimo del término "sistemas de IA de alto riesgo", utilizado en la propuesta de RIA de la UE. Ante estos riesgos, se evidencia la necesidad de construir conjuntamente "un enfoque regulatorio y de gobernanza proporcionado y a favor de la innovación que maximice los beneficios y tenga en cuenta los riesgos asociados con la IA, [...] en enfoques como principios y códigos de conducta comunes."³¹ Se observa cómo en esta declaración se reconoce el valor de las propuestas internacionales vigentes, que, son el RIA de la UE, , los los Principios Rectores Internacionales en materia de IA elaborado por la OCDE y el Código de Conducta del G-7. No obstante, se matiza en la propia Declaración que la elaboración de políticas basadas en el riesgo será conforme a los enfoques de cada territorio, que podrán diferir según las circunstancias y marcos jurídicos aplicables, lo que en gran medida puede difuminar la consecución de una norma única de carácter imperativo para todos los Estados y mercados.

5. CONCLUSIONES

1. La Unión Europea tiene la firme voluntad de liderar la regulación de la inteligencia artificial en el mundo. Al igual que ya sucedió en otros ámbitos, como la protección de los datos personales, la UE pretende construir un modelo normativo que sea replicado por el resto de los ordenamientos jurídicos extranjeros, lo que se conoce como el efecto Bruselas. No obstante, los objetivos propuestos en el RIA presentan un reto muy complejo, que genera un fundado escepticismo: incentiva la innovación y el desarrollo tecnológico en Europa, garantizando en todo momento que no se vulneren los derechos de los ciudadanos.

2. Fuera de las fronteras de la UE, también se está desarrollando una incesante actividad regulatoria en materia de IA. Como se observa entre las principales potencias tecnológicas, EE. UU. y China ya disponen de normas que imponen deberes de transparencia y control sobre el ciclo de vida de estos sistemas. Si bien, los esfuerzos aislados de los Estados por sí solos no son suficientes para regular esta tecnología, que tiene como característica principal la capacidad de influir en todos los ámbitos de la sociedad y en cualquier mercado nacional e internacional. Así, el Proceso de Hiroshima es el resultado de la búsqueda conjunta de unos estándares de fiabilidad en el desarrollo e implantación de los sistemas de la IA, aunque es una norma de mínimos insuficiente para solventar los riesgos que se presentan.

³¹ The Bletchley Declaration... op. cit. Ed. electr.

3. El enfoque europeo en torno a la IA y el Proceso de Hiroshima del G-7 muestran claras diferencias. La Unión Europea defiende un modelo con importantes garantías y salvaguardas, que ofrece una mayor seguridad jurídica, para consumidores y sujetos (proveedores, implantadores y representantes autorizados) responsables de los sistemas IA. A su vez, el Proceso de Hiroshima proyecta unas pautas básicas, en forma de principios y código de conducta, que obedece a la finalidad de construir un modelo común alineado con valores ampliamente aceptados como son los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos. Por tanto, pese a ser propuestas distintas sí son compatibles, pues parten de un punto común.

4. Por último, la celebración de la Cumbre de seguridad en Reino Unido tiene una finalidad distinta a la Cumbre del G-7 o los acuerdos producidos en el seno de la UE en el año 2023. Esta cumbre fomenta un diálogo mucho más amplio entre los Estados y el resto de agentes involucrados (representantes del sector tecnológico, expertos y académicos). En este sentido, más allá de medidas concretas por parte de los Estados, la cooperación es el verdadero instrumento imprescindible para aplacar los graves riesgos que ocasiona la IA.

6. BIBLIOGRAFÍA

B. FERNÁNDEZ, C. "Entra en vigor la Ley china sobre algoritmos de recomendación en Internet" *Derecho Digital e Innovación. Digital Law and Innovation Review*, núm. 13, 2022.

BARRIO ANDRÉS M. "Novedades en la tramitación del próximo Reglamento europeo de inteligencia artificial". En <https://www.realinstitutoelcano.org/analisis/novedades-en-la-tramitacion-del-proximo-reglamento-europeo-de-inteligencia-artificial/>.

BRUSSEE V., *Social credit. the warring states of China's emerging data empire*. Ed. el. Palgrave Macmillan. Singapore, 2023.

DAISUKE T. "New global norms on responsible AI beyond the EU: the G7 Hiroshima Process international guiding principles for developing advanced AI systems and its application in Japan". En: [https://www.ibanet.org/G7-Hiroshima-Process-international-guiding-principles-for-developing-advanced-AI-systems-application"-in-Japan#_ftn5](https://www.ibanet.org/G7-Hiroshima-Process-international-guiding-principles-for-developing-advanced-AI-systems-application).

JIMENO MUÑOZ, J. *Derecho de daños tecnológicos, ciberseguridad e insurtech*. 1 ed. Dykinson. Madrid, 2019.

TAPIA HERMIDA, A. "El Código de Conducta de la Inteligencia Artificial (CCIA) del G-7". En: <https://ajtapia.com/2023/11/el-codigo-de-conducta-de-la-inteligencia-artificial-ccia-del-g-7/>.

TAPIA HERMINDA, A. "Inteligencia Artificial responsable. Dos documentos relevantes para su regulación: El Informe Hiroshima sobre la Inteligencia Artificial Generativa y el Proyecto internacional de principios rectores para las organizaciones que desarrollan sistemas avanzados de Inteligencia Artificial". En: <https://ajtapia.com/2023/11/inteligencia-artificial-responsable-dos-documentos-relevantes-para-su-regulacion-el-informe-hiroshima-sobre-la-inteligencia-artificial-generativa-y-el-proyecto-internacional-de-principios-rectores-p/>.