

LA GESTIÓN DE RIESGOS COMO ENFOQUE REGULATORIO: CONSIDERACIONES SOBRE LA PROPUESTA DE REGLAMENTO SOBRE I.A. EN LA UNIÓN EUROPEA

RISK MANAGEMENT AS A REGULATORY APPROACH: SOME CONSIDERATIONS ON THE PROPOSAL FOR REGULATION OF IA IN THE EUROPEAN UNION

IRIT MILKES

RESUMEN:

La más reciente propuesta de Reglamento del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial, incorpora un enfoque normativo de gestión de riesgos. Esta contribución ofrece algunas consideraciones críticas alrededor de dicho enfoque, particularmente frente a comprensión de la futura ley de inteligencia artificial y los retos jurídicos que trae consigo para la evaluación de los riesgos derivados de sistemas de IA, las decisiones sobre su permisibilidad o prohibición y las actuaciones de control y seguimiento posteriores a su comercialización y/o implementación.

PALABRAS CLAVE: gestión de riesgos, regulación, control, sistemas de alto riesgo, análisis de riesgo.

ABSTRACT:

The most recent proposal from the European Parliament and the Council, establishing harmonized rules on artificial intelligence, incorporates a regulatory approach to risk management. This contribution offers some critical considerations around this approach, particularly regarding the understanding of the future artificial intelligence Act and the legal challenges it poses for the assessment of high-risk AI systems, decisions on their permissibility or prohibition, and subsequent monitoring and control actions after their commercialization or implementation.

KEYWORDS: risk management, regulation, monitoring and control, high-risk AI systems, risk assessment

SUMARIO: 1. Introducción: la gestión de riesgos como enfoque normativo en la propuesta de Reglamento Europeo del Parlamento y del Consejo en materia de Inteligencia Artificial. 2. La valoración cuantitativa y cualitativa de la permisibilidad del riesgo. La calificación de sistemas de IA de ‘alto riesgo’. 3. Medidas de gestión de riesgos y control poscomercialización de los sistemas de IA por agencias supervisoras. 4. Espacios de prueba controlados para la implementación de sistemas de IA de ‘alto riesgo’ por el sector público (*regulatory sandboxes*): una vía para la adecuación del interés general. 5. Bibliografía.

1. Introducción: la gestión de riesgos como enfoque normativo en la propuesta de Reglamento Europeo del Parlamento y del Consejo en materia de Inteligencia Artificial.

La más reciente propuesta de Reglamento del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial, incorpora un enfoque normativo de gestión de riesgos. En la redacción de la propuesta, se evidencia el interés en dotar de un tratamiento jurídico el riesgo que se deriva de la implementación de sistemas de inteligencia artificial, especialmente aquellos denominados como “sistemas de alto riesgo”. Muchas son las menciones que realizadas al ‘riesgo’ como categoría, su tratamiento y gestión, sin tener mayor claridad de lo que ello significa. El enfoque normativo en el que se basa la propuesta adquiere una relevancia significativa, en la medida que encuentra un espacio regulatorio de un fenómeno de constantes cambios y donde los instrumentos existentes se escabullen de los mecanismos tradicionales de regulación.

La propuesta tiene como fundamentos jurídicos, en primer lugar, el apartado primero del artículo 114 del Tratado de Funcionamiento de la Unión Europea (TFUE). En segundo lugar, el principio de subsidiariedad¹, en la medida que la inteligencia artificial depende de distintos conjuntos de datos que pueden ser tan variados como complejos y, por ende, pueden integrarse en otros productos y servicios que circulan libremente por el mercado interior. En tercer lugar, con el principio de proporcionalidad, se imponen cargas normativas solo en aquellos sistemas de inteligencia artificial que signifique un alto riesgo a la seguridad y protección de los ciudadanos, particularmente tras identificar este enfoque en todo el diseño de la propuesta².

El enfoque normativo en mención, busca principalmente subsanar, anticipar y gestionar los riesgos que puedan surgir de la implementación de un sistema de inteligencia artificial. Si bien la gestión de riesgos tradicionalmente se ha encargado de sectores estratégicos como la salud humana, el medioambiente o el cambio climático, todas estas reúnen un común denominador: la necesidad de garantizar la protección y seguridad de las personas frente a la introducción de nuevas tecnologías y la incertidumbre que se deriva de ellas³. En este sentido, la propuesta de Reglamento reconoce que la inteligencia artificial por su dinamismo, flexibilidad y constante cambio necesita de una regulación que se ajuste y adapte a ello. Evita el diseño de una regulación obsoleta y rígida que promueva la innovación, la creación de incentivos para su implementación y la generación de confianza y seguridad en los usuarios. El enfoque regulatorio basado en el riesgo segrega aquellos sistemas de I.A. que, tras su valoración y evaluación, no genera riesgos significativos y de

¹ Propuesta de Reglamento del Parlamento europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial, p. 6.

² Schuett, Jonas, Risk Management in the Artificial Intelligence Act, European Journal of Risk Regulation, 2023, pp. 1-19.

³ Esteve Pardo José, “El desconcierto del Leviatán. Política y Derecho ante la incertidumbre de la ciencia”, Marcial Pons, 2009, p. 211, p. 69.

generarlos, han sido aceptados socialmente como procesos inherentes a la evolución social y la adaptación de prácticas al surgimiento de nuevas tecnologías.

2. La valoración cuantitativa y cualitativa para la permisibilidad del riesgo. La calificación de sistemas de inteligencia artificial de ‘alto riesgo’.

El enfoque normativo de gestión de riesgos en la Propuesta de Reglamento, es el resultado de distintas evaluaciones sobre el riesgo derivado de sistemas de inteligencia artificial⁴. En el año 2019, la aprobación del Plan coordinado con los distintos Estados miembros contribuyó a la armonización de estrategias para su implementación en el ámbito europeo⁵.

La metodología de calificación de los distintos sistemas de IA tiene como presupuesto una evaluación multinivel del riesgo. En primer lugar, la decisión sobre la admisibilidad o prohibición de un riesgo, la antecede una evaluación realizada por el regulador y un grupo de expertos, quien determina -en primera instancia- cuáles de los sistemas serán prohibidos o permitidos. Tal es el caso de la evaluación realizada por un grupo de expertos de alto nivel sobre los riesgos permisibles de la I.A., donde se diseñaron directrices para alcanzar un nivel fiable y adecuado para la implementación de sistemas de inteligencia artificial⁶. La aceptación de su uso por los sistemas jurídicos ya no prioriza la prohibición de sistemas particulares, sino adopta una vía de autorizaciones condicionadas a garantizar la protección adecuada de los usuarios y un balance favorable en la relación de riesgo-beneficio⁷.

Una vez realizada la calificación, particularmente sobre aquellos señalados como ‘sistemas de alto riesgo’, la Comisión -en una segunda instancia- podrá modificar y agregar a la lista prevista (Art.6 y Anexo III) otros sistemas de IA que, según una valoración cuantitativa y cualitativa del riesgo, pueden potencialmente ocasionar algún daño. La valoración cuantitativa, analizará la gravedad y probabilidad del riesgo y, cuando superen los indicadores previstos, ésta determinará si modifica (o no) el listado de sistemas que implican un riesgo significativo o alto. Mientras que, la valoración cualitativa tendrá en cuenta criterios tales como la finalidad del sistema, las actualizaciones de las evaluaciones del riesgo anteriores en donde se demuestre las repercusiones negativas sobre las personas y el alcance de los perjuicios ocasionados a estas, e incluso, serán tenidas en cuenta las medidas de reacción para prevenir o reducir notablemente los riesgos, entre otros criterios (Art. 7). Por último, en una escala individualizada, cada proveedor de un sistema de

⁴ Desde el año 2017, se inició el proyecto normativo de armonizar las normas que regularían el uso de la inteligencia artificial en el espacio europeo, cuando el Consejo Europeo llegó a la conclusión que, con el objetivo de implementar el proyecto de Europa digital, se necesitaba de una regulación que cumpliera con este cometido.

⁵ COM 2018(795)

⁶ European Parliamentary Research Service. 2019. *EU Guidelines on Ethics in Artificial Intelligence: Context and Implementation*. Brussels: European Parliament Research Service.

⁷ Esteve Pardo, José (1999): *Técnica, Riesgo y Derecho. Tratamiento del riesgo tecnológico en el Derecho ambiental*, Ariel, Barcelona, p. 45. El autor explica el ejercicio de ponderación detrás de la toma de decisiones administrativas que busca promover beneficios y ventajas en entornos y sectores de riesgo.

IA de alto riesgo deberá realizar una evaluación de este a la hora de cumplir con el sistema de gestión de riesgos exigido en el Artículo 9.

La decisión sobre la permisibilidad de los riesgos derivados de algunos sistemas de IA, dependerá de su valoración previa. Así, aquellos que ostenten un riesgo excesivo, serían calificados como prohibidos (Art. 5), como es el caso de aquellos sistemas que se sirven de técnicas subliminales que inciden en el comportamiento de las personas; aquellos usados por autoridades públicas para la evaluación y clasificación de ciudadanos, atendiendo a sus conductas sociales o características de su personalidad; o incluso, el uso de sistemas de identificación biométrica remota ‘en tiempo real’ en espacios de acceso público, cuyo uso se permitiría estrictamente a efectos policiales y la prevención en la comisión de delitos graves.

En el mismo sentido tenemos que, algunos riesgos ‘limitados’ serían aceptados como parte inherente de los riesgos residuales por la apuesta en el plan de digitalización europea, a los cuales se les impondrían ligeras obligaciones de transparencia, como por ejemplo, la información a los destinatarios de decisiones públicas que hayan sido generadas por IA. Otros riesgos, si bien serían permitidos, se encontrarían bajo la condición de cumplir con algunas cargas en cabeza de los proveedores del sistema. Esta última categoría de riesgo, calificada como “sistemas de alto riesgo” en la propuesta de Reglamento (Art.6), si bien representan un riesgo significativo para la protección de derechos fundamentales o la seguridad y salud de las personas, son aceptados bajo la condición de cumplir, obligatoriamente, requerimientos técnicos que serán objeto de supervisión, tanto de la autoridad europea como de las autoridades competentes en cada uno de los Estados miembros. Puede ser el caso particular de los sistemas que figuran en el Anexo III, en el que se encuentran sistemas de IA que tienen por finalidad: la identificación biométrica y categorización de personas físicas, la gestión y funcionamiento del tráfico, el suministro de servicios públicos esenciales, la asignación o la determinación del acceso a centros de educación, la contratación o selección de personas físicas para empleo, el acceso a servicios públicos y privados esenciales y sus beneficios, la determinación y valoración del riesgo de infracción y comisión de delitos personas físicas, la gestión migratoria y el control fronterizo, entre otros.

La propuesta de Reglamento no ofrece detalles específicos respecto a su alcance sobre las Administraciones públicas. La interpretación de éste sugiere que, las entidades del sector público, ya sea en calidad de proveedoras o usuarias de sistemas de inteligencia artificial, también estarán sujetas a este nuevo marco regulatorio. A pesar de que el futuro Reglamento establecerá un marco sobre los riesgos prohibidos u aceptados, serán las autoridades nacionales de supervisión de cada Estado miembro las encargadas de adaptar la propuesta normativa a las entidades del sector público. Esta adaptación, sin lugar a dudas, deberá abordar las cuestiones particulares que deben incorporarse en el diseño de un estatuto jurídico específico para las entidades de la Administración pública y las necesidades propias que estas plantean.

En los términos expuestos en la propuesta, gran parte de las actuaciones públicas podrían resultar en la clasificación de sistemas como 'alto riesgo', sin que estas impliquen una amenaza significativa para la seguridad o salud de las personas. En consecuencia, será fundamental cuestionar si los criterios de evaluación de riesgos para los sistemas de inteligencia artificial, como se detallan en el futuro Reglamento, son aplicables o no al sector público. Consideramos que se deben llevar a cabo evaluaciones específicas que contemplen aspectos como la estructura organizativa de las entidades públicas, el alcance de las competencias asignadas, y la viabilidad técnica o financiera para cumplir con los requisitos establecidos. Estos elementos deben formar parte integral de la evaluación cualitativa de los riesgos derivados de los sistemas de inteligencia artificial en las instituciones de la Administración pública⁸.

3. Medidas de gestión de riesgos y control poscomercialización por las Agencias supervisoras.

Los sistemas de gestión de riesgo, como requisito obligatorio para proveedores de sistemas de IA de alto riesgo, ejemplifique el enfoque normativo propuesto por la Comisión. Se trata de un proceso iterativo llevado a cabo durante todo el ciclo de vida de los sistemas de IA, donde, después que son permitidos por la ley o el reglamento, cada proveedor continúa realizando evaluaciones individualizadas y periódicas. El objetivo es determinar medidas oportunas y apropiadas para eliminar o reducir los riesgos derivados del sistema, así como los riesgos residuales asociados con el uso razonablemente previsible de estos, como se establece en los apartados 3 y 4 del Art. 9.

Considerando la imposibilidad, en la mayoría de los casos y a menos que el estado de la técnica demuestre lo contrario, de alcanzar un escenario de 'peligro cero', se implementarán medidas de reducción, mitigación y control adecuadas para aquellos riesgos que, de hecho, no puedan ser eliminados. Antes de su introducción en el mercado o puesta en servicio, los sistemas clasificados como 'alto riesgo' deberán someterse a procedimientos de prueba para asegurar la seguridad de los usuarios y evaluar los riesgos asociados con su comercialización. Estos ensayos constituirán una parte integral de la evaluación cuantitativa y cualitativa del riesgo a escala individual. En este contexto, el proveedor, basándose en parámetros y umbrales de probabilidades previamente definidos, determinará si el producto es adecuado para alcanzar la finalidad prevista.

Dado que se trata de un proceso iterativo, las evaluaciones periódicas son esenciales para garantizar la seguridad y salud de los usuarios de sistemas considerados de 'alto riesgo', como por ejemplo puede ser la exigencia adicional de la vigilancia humana. Esto implica que, al diseñar un sistema de inteligencia artificial, se debe garantizar que pueda ser supervisado por personas físicas durante su uso, además de dotarlo de herramientas de interfaz humano-máquina. En este sentido, la

⁸ Algún sector de la doctrina ha enfatizado en el valor diferencial del diseño normativo sobre IA en el ámbito del sector público. En este sentido ver: Cerrillo i Martínez, Agustí. 2020c. El Impacto de La Inteligencia Artificial En Las Administraciones Públicas: Estado de La Cuestión y Una Agenda. In *Retos Jurídicos de La Inteligencia Artificial*. Edited by Agustí Cerrillo i Martínez and Miquel Peguera Poch. Pamplona: Thomson Reuters Aranzadi, pp. 75–92.

vigilancia llevada a cabo por una persona física, aunque no excluye el control externo que puedan llevar a cabo las autoridades competentes en la supervisión de la inteligencia artificial, permite observar y detectar cualquier indicio de anomalías, problemas de funcionamiento o comportamientos inesperados y, posibilita adoptar una reacción oportuna ante cualquier eventualidad (Apartados 4 del Artículo 14 y del Artículo 21).

Adicionalmente, el sistema de gestión de riesgos y las medidas propuestas para su tratamiento en una escala individual permitirá: primero, entender las capacidades y limitaciones del sistema de IA; segundo, evitar los sesgos de automatización e interpretar de forma correcta las salidas del sistema de IA, más aún cuando las Administraciones públicas, en calidad de proveedores o usuarias, toman decisiones públicas apoyadas en estos; y tercero, decidir la no utilización de sistemas de IA de alto riesgo y desestimar las salidas que este proporcione cuando de su desarrollo se provoque o generen potenciales daños a la salud o seguridad de las personas.

Los planes de seguimiento posteriores a la comercialización (Art. 61) serán parte de las actuaciones de autocontrol por los proveedores de sistemas de IA que, de igual forma, contribuirán con el ejercicio de actividades de supervisión de las autoridades nacionales y comunitarias competentes. En la misma línea, los requerimientos de información a proveedores o usuarios, notificaciones de incidencias o problemas del sistema a la autoridad competente serán necesarias para la ejecución de medidas de gestión del riesgo y su seguimiento por la Administración pública en posición de garante (Art. 29).

En lo que respecta a la gobernanza de la inteligencia artificial y la supervisión de su uso tanto en el sector público como en el privado, se contempla la creación de un Comité Europeo de Inteligencia Artificial (Artículo 56). Este comité desempeñará un papel crucial a la hora de ofrecer asesoramiento y asistencia a la Comisión, asegurando la colaboración en el ejercicio de diversas funciones relacionadas con el control jurídico y técnico, la armonización de las prácticas administrativas entre los Estados miembros, la orientación y guía en la materia, así como la realización de actividades conjuntas con las autoridades nacionales de supervisión (Artículo 58).

Por último, cada Estado miembro designará una Agencia Nacional de Supervisión, la cual, en el caso español, ya ha sido establecida mediante la expedición de la Ley 22/2021 de Presupuestos Generales del Estado, para el año 2022. La Agencia Española de Supervisión de la Inteligencia Artificial (AESIA) actuará como autoridad notificante y entidad vigilante de acuerdo con lo establecido en el Artículo 59.

4. Los espacios de prueba controlados para la implementación de sistemas de IA de ‘alto riesgo’ por el sector público: una vía para la adecuación de intereses generales.

Con la implementación de un nuevo marco normativo en materia de inteligencia artificial, particularmente a lo que al sector público se refiere, los entornos controlados de pruebas permitirán incentivar su uso por entidades públicas y, además, garantizar los mecanismos para adaptar su

cumplimiento a las necesidades específicas de éstas. Los esfuerzos para el cumplimiento normativo que exige la regulación del desarrollo tecnológico, necesita -en ocasiones- entornos de experimentación supervisados por las autoridades de control. Por tal razón, esta posibilidad significa la apertura de una vía para la adaptación de sistemas de IA a los intereses generales perseguidos por la Administración pública, garantizar la eficiencia de la actividad administrativa⁹, así como también la materialización de principios como el de ‘buena administración’¹⁰.

Conscientes de los esfuerzos, tanto técnicos como financieros de la implementación de sistemas de IA por el sector público, la posibilidad de analizar la evolución y posibles impactos, así como el diseño y adaptación de estos y su alineación al interés general perseguido por las Administraciones públicas permitiría ofrecer un enfoque proporcionado y adecuado a las exigencias formales y materiales que exige la actuación administrativa. En España, las experiencias ya integradas en el sistema financiero, por la Ley 7/2020, de 13 de noviembre para la transformación digital del sistema financiero, y del sector eléctrico y, por el Real Decreto 568/2022, de 11 de julio, por el que se establece el marco general del banco de pruebas regulatorio para el fomento de la investigación y la innovación en el sector eléctrico, sirven de base para su extensión en distintos ámbitos del sector público.

El reciente Real Decreto 817/2023, de 8 de noviembre, que establece un entorno controlado de pruebas para evaluar el cumplimiento de la propuesta de Reglamento sobre inteligencia artificial, define los ámbitos de aplicación. En este entorno de prueba, se incluyen como participantes a las Administraciones públicas y entidades del sector público que utilicen sistemas de IA de 'alto riesgo' y sistemas de propósito general. Además, se contempla la incorporación de sistemas más innovadores, como los modelos fundacionales para la buena administración. No obstante, se excluye de la participación en este entorno a los sistemas de IA cuyos propósitos sean gestionar actividades militares, de defensa o seguridad nacional, utilizar técnicas subliminales para alterar el comportamiento de las personas físicas, aprovecharse de vulnerabilidades de un grupo específico, y otras acciones ya prohibidas según la propuesta de Reglamento

Las expectativas de su implementación no solo corresponden a la adaptación y adecuación de instrumentos tecnológicos al ejercicio de actividades públicas, sino también la elaboración de guías técnicas de ejecución y supervisión de la AESIA e informes de buenas prácticas específicas para el sector público. En este caso, si bien no se persigue la elaboración de un marco regulatorio, objetivo principal de los espacios controlados de prueba, se pretenderá la adaptación de un marco normativo y la consolidación de la confianza en los sistemas de IA.

A pesar que, las disposiciones normativas del Real Decreto señalan que podrán ser objeto de espacios controlados de prueba los sistemas que ya estén comercializados o cuyo desarrollo se

⁹ Zlotnik, Alexander. 2019. Inteligencia Artificial En Las Administraciones Públicas: Definiciones, Evaluación de Viabilidad de Proyectos y Áreas de Aplicación. *Boletín* 84: 25–32

¹⁰ Ponce Solé, Juli. 2018. La Prevención de Riesgos de Mala Administración y Corrupción, La Inteligencia Artificial y El Derecho a Una Buena Administración. *Revista Internacional de Transparencia e Integridad* 6: 1–19.

encuentra en una fase de cambios sustanciales que facilite la introducción de medidas necesarias para cumplir con el futuro Reglamento, estos escenarios de experimentación controlada serán esenciales para la transición digital del sector público y la orientación suficiente para capacitar a aquellos encargados de las entidades públicas en su adaptación.

Por último, el enfoque normativo de la gestión de riesgos tanto en el diseño normativo, como en los instrumentos jurídicos que facilitan su adaptación, tienen una finalidad común: la reducción de riesgos derivados del uso de la inteligencia artificial y la mitigación de los riesgos residuales que pudiesen concurrir para la materialización de daños y perjuicios en la seguridad y la salud de las personas.

BIBLIOGRAFIA

- Cerrillo i Martínez, Agustí (2020): “El Impacto de La Inteligencia Artificial En Las Administraciones Públicas: Estado de La Cuestión y Una Agenda” en A. Cerrillo i Martínez y M. Peguera Poch, *Retos Jurídicos de La Inteligencia Artificial*. Pamplona, Thomson Reuters Aranzadi, pp. 75–92.
- Esteve Pardo, José (2009): *El desconcierto del Leviatán. Política y Derecho ante la incertidumbre de la ciencia*, Marcial Pons, p. 69.
- Esteve Pardo, José (2015): “La administración garante. Una aproximación”, en *Revista de Administración Pública*, num. 197, mayo-agosto, pp. 11-39.
- Esteve Pardo, José (1999): *Técnica, Riesgo y Derecho. Tratamiento del riesgo tecnológico en el Derecho ambiental*, Ariel, Barcelona, pp. 20-56.
- European Parliamentary Research Service. 2019. *EU Guidelines on Ethics in Artificial Intelligence: Context and Implementation*. Brussels: European Parliament Research Service. Disponible en: <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>
- Huergo Lora, Alejandro (2021): El Proyecto de Reglamento Sobre La Inteligencia Artificial. *Almacén de Derecho*. Disponible en: <https://almacenederecho.org/el-proyecto-de-reglamento-sobre-la-inteligencia-artificial>
- Ponce Solé, Juli (2018): “La Prevención de Riesgos de Mala Administración y Corrupción, La Inteligencia Artificial y El Derecho a Una Buena Administración”, en *Revista Internacional de Transparencia e Integridad*, núm. 6, pp. 1–19.
- Schuett, Jonas (2023): “Risk Management in the Artificial Intelligence Act” en *European Journal of Risk Regulation*, num. 9, pp. 1-19.
- Van Roy, Vicent, Fiammetta Rossetti, Karine Perset, and Laura Galindo-Romero (2021): *AI Watch—National Strategies on Artificial Intelligence: A European Perspective*, 2021th ed. JRC Working Papers No. JRC122684. Brussels: Publications Office of the European Union
- Zlotnik, Alexander (2019): “Inteligencia Artificial En Las Administraciones Pública: Definiciones, Evaluación de Viabilidad de Proyectos y Áreas de Aplicación” en *Boletín*, núm. 84, pp. 25–32