Real Decreto 817/2023, de 8 de noviembre, 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 [BOE-A-2023-22767]

ENTORNO CONTROLADO DE PRUEBAS PARA UNA INTELIGENCIA ARTIFICIAL CONFIABLE

1. CONSIDERACIONES GENERALES

Teniendo en cuenta el enorme crecimiento de la innovación basada en la inteligencia artificial (en adelante IA), se hace cada vez más necesario establecer una regulación sólida que garantice su uso ético y responsable. Con este objetivo, la Comisión Europea ha presentado una propuesta de Reglamento del Parlamento Europeo y del Consejo (también conocido como futura Ley europea de Inteligencia Artificial) por el que se establecen normas armonizadoras en materia de inteligencia artificial, con el propósito de asegurar el respeto de los derechos fundamentales de la ciudadanía y generar confianza en el desarrollo y la utilización de la inteligencia artificial de manera holística en la economía y la sociedad; creando un mercado único digital para la IA y facilitando su adopción en todos los sectores y actividades sociales.

El citado reglamento busca proveer a la Unión Europea de un marco normativo con el fin de promover una inteligencia artificial fiable, ética y robusta, centrándose, más que en la tecnología en sí, en las aplicaciones de la IA, cuyo potencial incide positivamente en el crecimiento económico, la creación de empleo y el progreso social. Siendo conscientes, a su vez, de que los sistemas de IA también pueden suponer riesgos sobre el respeto de los derechos fundamentales (como los relativos a la discriminación y a la protección de datos personales) o incluso causar problemas graves sobre la salud o la seguridad de la ciudadanía.

Por todo ello, a la espera de la aprobación del Reglamento de la UE, el Gobierno de España, con la colaboración de la Comisión Europea, pone en marcha el primer entorno controlado de pruebas *(sandbox)* para ensayar la aplicación de ciertos requisitos previstos en la propuesta de Reglamento de Inteligencia Artificial a los sistemas de inteligencia artificial de alto riesgo, con el objetivo de obtener, como resultado de este ensayo, unas guías basadas en la evidencia y la experimentación que ayuden a las empresas y a la sociedad en general al cumplimiento de la propuesta de Reglamento de IA.

La finalidad de estas pruebas es la de llevar a cabo una autoevaluación de las aplicaciones de IA, encuadrando todo ello en el plan de digitalización de la Agenda 2026 y en el mismo Plan de Recuperación, Transformación, y Resiliencia, cuyo componente

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

16 tiene como objetivo apoyar el despliegue y el uso masivo de la inteligencia artificial por parte de las grandes empresas, las Administraciones Públicas, las pequeñas y medianas empresas y empresas emergentes y la sociedad civil.

El RD 817/2023 se dicta de conformidad con «la habilitación prevista en el artículo 16 de la Ley 28/2022, de 21 de diciembre, de Fomento del Ecosistema de las Empresas Emergentes, donde se contempla la creación de entornos controlados, por períodos limitados de tiempo, para evaluar la utilidad, la viabilidad y el impacto de innovaciones tecnológicas aplicadas a actividades reguladas, a la oferta o provisión de nuevos bienes o servicios, a nuevas formas de provisión o prestación de los mismos o a fórmulas alternativas para su supervisión y control por parte de las autoridades competentes». Esta misma norma señala que «la creación de los entornos controlados de pruebas para la evaluación de su impacto, está justificada por razones imperiosas de interés general». Cabe destacar que, inspirada por la Carta de Derechos Digitales, esta iniciativa pretende dar una forma concreta y práctica al compromiso español de «establecer un marco ético y normativo que refuerce la protección de los derechos individuales y colectivos».

En definitiva, el RD 817/2023 pretende establecer un entorno controlado de pruebas para ensayar el cumplimiento de ciertos requisitos por parte de algunos sistemas de IA que puedan suponer *riesgos para la seguridad, la salud y los derechos fundamentales de las personas*. Para ello, se regula el procedimiento de selección de los sistemas y entidades que participarán en el entorno controlado de pruebas (art. 1), siendo de aplicación tanto a las administraciones públicas y entidades del sector público institucional, tal y como se define en el artículo 3.14, como a entidades privadas seleccionadas en el entorno controlado de pruebas de IA. El objeto de este entorno será estudiar la operatividad de los requisitos establecidos en la propuesta de reglamento europeo, la realización de una autoevaluación de cumplimiento de los mismos y la evaluación del plan posterior a la comercialización de los sistemas de IA de las entidades participantes.

Por lo que se refiere a la *estructura del RD 817/2023* (de vigencia limitada, como reza su disposición adicional segunda), consta de un título preliminar, dos títulos, dos disposiciones adicionales, dos disposiciones finales y siete anexos. El título preliminar contiene las disposiciones generales, estableciendo su objeto y ámbito de aplicación, así como las definiciones de los conceptos principales a efectos de lo previsto en la norma. El título I se estructura en seis capítulos, a través de los cuales se concretan los requisitos para la elegibilidad y la participación en el entorno controlado de pruebas, para lo que se regula el régimen jurídico aplicable; la figura del proveedor de sistema de inteligencia artificial; los criterios de elegibilidad; el modo de participación y procedimiento de admisión; la forma en que se evaluarán las solicitudes; las particularidades y condiciones concretas para el desarrollo de esta experiencia, y las garantías de las entidades participantes; los canales de comunicación y la finalización de la experiencia en el entorno. Por su parte, el título II contiene una serie de disposiciones relativas a la colaboración y coordinación entre autoridades, personas asesoras expertas y otros organismos españoles y europeos. Las dos disposiciones adicionales se ocupan: de los

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

medios a disposición del entorno controlado de pruebas, que serán los de la Secretaría de Estado de Digitalización e Inteligencia Artificial; del resultado del entorno controlado de pruebas, que deberá ser publicado en su portal web junto con el informe de las conclusiones sobre el desarrollo, buenas prácticas y recomendaciones al mismo, así como otros aspectos de interés; además, el órgano competente (Secretaría de Estado de Digitalización e Inteligencia Artificial), partiendo de los resultados obtenidos en el entorno controlado de pruebas, podrá desarrollar una plataforma de software que facilite una primera autoevaluación no vinculante sobre el cumplimiento de los principios de la propuesta del Reglamento de la UE. Y, en sus dos disposiciones finales se indica el título competencial, la vigencia y la entrada en vigor. Finalmente, decir que los siete anexos que lo conforman son de extraordinaria importancia, puesto que se convierten en «normas técnicas» al establecer los criterios prácticos de aplicación del mismo. Estos anexos se pronuncian sobre las siguientes cuestiones:

- El primero da cuenta de la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial.
- El segundo contiene un listado de áreas de sistemas de inteligencia artificial de alto riesgo específicos, como los sistemas de identificación biométrica o los que se relacionen con el sistema educativo, laboral, con la aplicación de la ley o con la actividad jurisdiccional, u otros especialmente sensibles.
- El tercero se pronuncia sobre el contenido mínimo de la memoria técnica para la solicitud de participación en el entorno de pruebas.
- El cuarto versa sobre la declaración responsable de cumplimiento del principio de responsabilidad proactiva en materia de protección de datos.
- El quinto sobre documentación que se podrá requerir para el cumplimiento de la normativa del tratamiento de datos de carácter personal.
- El sexto sobre documentación técnica a presentar a la finalización de la implantación de los requisitos.
- Y el séptimo incluye un listado de legislación de la Unión Europea basada en el nuevo marco legislativo.

Hecho este planteamiento general acerca del contenido de la norma, vamos a comentar algunas cuestiones que, consideramos, suscitan especial interés.

2. SISTEMAS DE INTELIGENCIA ARTIFICIAL PARTICIPANTES EN EL ENTORNO CONTROLADO DE PRUEBAS Y NIVELES DE RIESGO

El RD 817/2023 define el «Sistema de inteligencia artificial» como un

> sistema diseñado para funcionar con un cierto nivel de autonomía y que, basándose en datos de entradas proporcionadas por máquinas o por personas, infiere cómo lograr

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

> un conjunto de objetivos establecidos utilizando estrategias de aprendizaje automático o basadas en la lógica y el conocimiento, y genera información de salida, como contenidos (sistemas de inteligencia artificial generativos), predicciones, recomendaciones o decisiones, que influyan en los entornos con los que interactúa. (art. 3.3)

Pues bien, según el nivel de riesgo para el usuario y la sociedad, el RD 817/2023 contempla dos clases de sistemas de IA con distinto alcance y efectos: sistemas de inteligencia artificial de alto riesgo (art. 3.4 y Anexo II) y sistemas de inteligencia artificial de propósito general (art. 3.5).

a) **Se entiende por sistemas de alto riesgo**:

- Los que constituyan productos regulados por la legislación de armonización de la Unión Europea en diversos ámbitos (máquinas, juguetes, embarcaciones de recreo y motos acuáticas, ascensores, aparatos y sistemas de protección para uso en atmósferas potencialmente explosivas, equipos radioeléctricos, equipos a presión, instalación de transportes por cable, equipos de protección individual, aparatos de quema de combustibles gaseosos y productos sanitarios), siempre que se exija que se sometan a evaluación de conformidad por terceros con vistas a su introducción en el mercado o puesta en servicio.
- Los que vayan a ser utilizados como componentes de seguridad de un producto regulado por una norma de armonización de la Unión Europea, en los mismos casos que los anteriores.
- Y determinados sistemas de inteligencia artificial relativos a la biometría, las infraestructuras críticas, la educación y la formación profesional, el empleo, el acceso a servicios públicos y privados, la persecución de delitos, la gestión de la migración, el asilo y el control fronterizo y la administración de justicia, en los casos en los que incidan en acciones o decisiones a tomar y puedan provocar un riesgo significativo para la salud, la seguridad o los derechos fundamentales.

b) **Por lo que se refiere a los sistemas de propósito general**, son definidos como los destinados por su proveedor a realizar funciones de aplicación general, como el reconocimiento de textos e imágenes, generación de audios y vídeos, respuestas a preguntas, traducciones, etc. En este grupo se incluyen los *modelos fundacionales* que el real decreto no define, pero que son aquellos sistemas entrenados y diseñados para adaptarse a una amplia gama de tareas posteriores; y los *sistemas de inteligencia artificial generativa*, que son modelos fundacionales destinados específicamente a generar, con distintos niveles de autonomía, contenidos tales como textos complejos, imágenes, audios o vídeos.

Por todo ello, atendiendo a los distintos niveles de riesgo para el usuario y la sociedad, la propuesta de Reglamento clasifica los sistemas de IA en tres categorías: *sistemas prohibidos*, *de riesgo limitado* y de *riesgo mínimo*.

Los sistemas prohibidos son aquellos que contravienen los valores europeos o causan daños inaceptables a las personas o a sus derechos. Estos sistemas incluyen

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

aquellos que manipulan el comportamiento humano o explotan las vulnerabilidades, los que permitan una vigilancia social masiva e indiscriminada o aquellos que evalúan aspectos sociales como el crédito o el comportamiento.

Frente a los anteriores, los sistemas de riesgo limitado implican una comunicación entre el usuario humano y la IA, pero no tienen un impacto significativo sobre ellos, debiendo informar al usuario de manera clara y fehaciente cuando esté interactuando con una máquina. Por último, los sistemas con riesgo mínimo son aquellos que tienen un impacto insignificante o nulo sobre las personas o sus derechos, debiendo cumplir con las normas generales aplicables.

Sin perjuicio de la categoría a la que respondan esos sistemas de IA, *el real decreto excluye del entorno de pruebas los sistemas de IA destinados exclusivamente* a fines de investigación (salvo para determinados supuestos específicos de investigación de delitos o prevención de amenazas) y a fines científicos; los que tengan una finalidad militar y de defensa o de seguridad nacional; aquellos que se sirvan de técnicas subliminales para alterar el comportamiento de las personas cuando puedan provocarles daños; los que se aprovechan de vulnerabilidades; los destinados a realizar clasificaciones sociales de personas que den lugar a tratos perjudiciales; o, sin perjuicio de algunas excepciones, los sistemas de identificación biométrica en tiempo real (los datos biométricos son datos personales relativos a las características únicas del ser humano, sean físicas, fisiológicas o asociadas al comportamiento, que facilitan y garantizan la identificación de un individuo [persona física], mediante sistemas o procedimientos tecnológicos —por ej. huella dactilares o la voz—).

Por otro lado, manifestar que la participación en el *sandbox* estará abierta a las siguientes personas:

- **Proveedores** de sistemas de inteligencia artificial, entendiendo por estos toda persona jurídica privada, Administración Pública, entidad del sector público u organismo de otra índole que haya desarrollado o para quien un tercero haya desarrollado un sistema de inteligencia artificial, cuando lo introduzca en el mercado o lo ponga en servicio bajo su propio nombre o marca comercial en calidad de proveedor. Los proveedores podrán presentar uno o varios sistemas diferentes, pero solo podrán ser admitidos a participar en relación con uno de ellos.
- Los **usuarios** residentes o establecidos en España, definidos como las personas jurídicas (no así las físicas), Administraciones Públicas o entidades del sector público bajo cuya autoridad se utilice un sistema de inteligencia artificial. Los usuarios solo podrán acceder al *sandbox* si el proveedor correspondiente también lo hace.

La participación de proveedores y usuarios en el entorno controlado de pruebas, que siempre será voluntaria (y con libre retirada), requerirá efectuar una solicitud formal una vez que por resolución de la autoridad competente se publique la convocatoria correspondiente y se especifiquen los requisitos exigibles. Inicialmente, el real decreto atribuye esa función a la Secretaría de Estado de Digitalización e Inteligencia Artificial, si bien, la reciente aprobación del Estatuto de la Agencia Española de Supervisión de

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

Inteligencia Artificial por el Real Decreto 729/2023, de 22 de agosto, podría llevar a que finalmente se adoptara una solución distinta.

3. PROTECCIÓN DE DATOS PERSONALES Y OBSERVANCIA DE LOS DERECHOS DE PROPIEDAD INTELECTUAL

Uno de los temas que más preocupan es la protección de los datos, ya que a través de estos sistemas de IA se recopilan ingentes cantidades de datos que, a su vez, pueden contener información sensible o personal. De ahí que, para promover el desarrollo y el uso responsable, sostenible y confiable de la IA, se ha creado, entre otros, la Agencia Española de Supervisión de la Inteligencia Artificial, convirtiéndose España en el primer país de la Unión Europea en dotarse de una autoridad de este tipo, que será clave para asegurar que se cumple con lo establecido en la futura ley de IA, así como para supervisar la correcta interpretación de la misma.

De esta forma, en las propuestas presentadas en el marco del entorno controlado de pruebas, *deben definirse de forma clara las posiciones jurídicas de cada uno de los intervinientes respecto de los tratamientos de datos personales*, ya sea bien como responsables, corresponsables, encargados o subencargados del tratamiento, conforme establece el Reglamento (UE) 679/2016 (RGPD), cuya observancia no es desplazada, sino reforzada por el Real Decreto 817/2023, como así se desprende de su art. 16, de forma que tanto los proveedores de IA participantes como los usuarios participantes en el entorno controlado de pruebas respetarán las disposiciones de protección de datos aplicables. *El régimen de protección de datos de carácter personal en las actuaciones que se desarrollen en el marco de este entorno controlado de pruebas es el previsto tanto en el* Reglamento (UE) 679/2016, del Parlamento Europeo y el 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 (RGPD), como en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales (LOPDGDD), y, en lo que resulte de aplicación, en la Ley Orgánica 7/2021, de 26 mayo, de protección de datos personales, debiendo quedar todos los datos personales que se traten en los sistemas privados de los proveedores de IA participantes o, en su caso, de los usuarios participantes, según corresponda, para comprobar los requisitos incluidos bajo dicha protección. Ello sin perjuicio de que alguno de los sistemas de alto riesgo indicados en el Anexo II del RD, que puedan participar en el entorno controlado de pruebas, deban tener un análisis previo de la licitud del tratamiento.

Por tanto, la aceptación de participación en el entorno controlado de pruebas implicará reconocimiento del cumplimiento de la legislación en materia de protección de datos. Por su parte, el Anexo IV del RD establece la obligación de presentar una declaración responsable con relación al cumplimiento de la normativa relativa a la protección

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

de datos personales y de aquella documentación que se estableciera en la convocatoria (Anexo V del RD).

Asimismo, los proveedores de IA y los usuarios participantes en el entorno controlado de pruebas deberán cumplir con lo previsto en la normativa de propiedad intelectual. La aceptación de la participación en el entorno controlado de pruebas implicará el compromiso y el reconocimiento del cumplimiento de la normativa en materia de propiedad intelectual. En esta línea, la futura ley europea de IA establecerá requisitos de transparencia, imponiendo, tanto a las máquinas como a sus usuarios, la obligación de operar en conformidad con la legislación de derechos de autor vigente en la Unión Europea. De igual forma, se exigirá la publicación de resúmenes explicativos del contenido utilizado para entrenar a la inteligencia artificial junto con la obra correspondiente.

4. RESPONSABILIDAD DE LOS PARTICIPANTES

Otra de las cuestiones que suscitan especial interés es la referida al *régimen de responsabilidad de los participantes*. Tanto el proveedor IA participante como, en su caso, el usuario participante serán responsables de los daños sufridos por cualquier persona como consecuencia de la aplicación del sistema de inteligencia artificial en el contexto del entorno controlado de pruebas, siempre que dichos daños deriven de un incumplimiento o cuando medie culpa, negligencia o dolo por su parte (art. 17.1 RD).

A este respecto, desde las instituciones europeas se han adoptado algunas propuestas, así, el Parlamento Europeo dictó una Resolución, de fecha 16 de febrero de 2017, que contiene recomendaciones destinadas a la Comisión sobre normas de Derecho Civil en materia de robótica.

Esta resolución se hace eco del debate acerca de si la normativa general sobre responsabilidad civil es suficiente o si, por el contrario, se requieren normas y principios específicos que aporten claridad sobre las distintas responsabilidades de los agentes intervinientes, argumentando que, en el actual marco jurídico, los robots no pueden ser responsables de los actos u omisiones que causan daños a terceros, en tanto en cuanto las normas vigentes en materia de responsabilidad civil contemplan los casos en los que únicamente es posible atribuir la acción u omisión del robot a un agente humano concreto, ya sea el fabricante, el operador, el propietario o el usuario.

A su vez, en materia de responsabilidad, recomienda establecer un *régimen de seguro obligatorio*, que los fabricantes y propietarios de los robots estarían obligados a suscribir por los posibles daños y perjuicios causados por estos, abogando por la creación de un fondo de compensación para los casos en los que no se haya contratado el seguro. Esta medida permite que el fabricante, programador, propietario o usuario del robot pueda beneficiarse de un sistema de responsabilidad limitada, siempre y cuando este contribuya al fondo de compensación o bien suscriba conjuntamente un seguro que garantice la compensación de daños y perjuicios causados por los robots.

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA

Finalmente, la resolución plantea crear a largo plazo una personalidad jurídica específica para robots, de forma que al menos los robots autónomos más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que puedan causar, aplicando esa personalidad electrónica a aquellos supuestos en los que los robots tomen decisiones autónomas inteligentes o sean capaces de interactuar con terceros de forma independiente.

Por su parte, la Resolución de 2020 del Parlamento Europeo incluye *recomendaciones a la Comisión sobre el régimen de responsabilidad civil en materia de IA*. Esta resolución supone un nuevo paso en el proceso de regulación de la robótica y la inteligencia artificial. Es de las soluciones más novedosas en este ámbito a nivel mundial con su propuesta de un reglamento sobre IA que incluye una normativa especializada sobre la responsabilidad civil cuando el daño sea ocasionado por algún sistema inteligente.

En la propuesta de Reglamento, contenida en la citada resolución, el Parlamento parte de que la persona que cree, mantenga, controle, explote el sistema de IA ha de ser responsable del daño o perjuicio que cause el dispositivo o la actividad que lleve a cabo el mismo. Partiendo de esta premisa, el Parlamento entiende que la Directiva 85/374 del Consejo sobre responsabilidad por los daños por productos defectuosos puede aplicarse en relación con las reclamaciones por responsabilidad civil formuladas frente al productor de un sistema de IA defectuoso. La responsabilidad civil del operador se basa en el hecho de que este ejerce un grado de control sobre un riesgo asociado al funcionamiento y la operación de un sistema de IA. Ahora bien, el riesgo predicable de los distintos sistemas de IA no es equivalente, así, para los sistemas de IA de alto riesgo (aquellos que funcionan de forma autónoma), la propuesta de Reglamento gira sobre un sistema de responsabilidad objetiva del operador, tal y como se recoge en el artículo 4. En cambio, el operador de un sistema de IA que no sea de alto riesgo estará sujeto a un régimen de responsabilidad subjetiva (artículo 8). En este último caso, el operador no será responsable cuando pueda demostrar que no tuvo culpa o negligencia en el daño causado.

María José CALVO SAN JOSÉ
Profesora de Derecho Civil
Universidad de Salamanca
calvo@usal.es

Ars Iuris Salmanticensis,
vol. 12, Junio 2024, 231-238
eISSN: 2340-5155
Ediciones Universidad de Salamanca - CC BY-NC-SA