

# CARTA DE DERECHOS DIGITALES

## JUSTIFICACIÓN GENERAL

El desarrollo de los sistemas de Inteligencia Artificial (IA) ha tenido un gran impulso en los últimos años y ha pasado a integrarse en gran parte de las aplicaciones de nuestra vida diaria. Por ello, una correcta regulación de los mismos resulta clave para la protección de los ciudadanos que hacen uso de estos sistemas.

Esta regulación comienza de algún modo en España con la Carta de Derechos Digitales que ahora se presenta. Hasta el momento se encuentra previsto en nuestra normativa, a distinto nivel, la habilitación para el uso de procesos que integran sistemas de IA; pero probablemente esta sea la base sobre la que se asienten los principios de posteriores normas sobre IA ética y segura que quizás solo se habían tratado antes a nivel de protección de datos.

En este asentamiento de las bases, la empresa desea aportar unos breves comentarios a tener en cuenta en relación al TÍTULO XXIII.

Los sistemas que incorporan Inteligencia Artificial pueden ser muchos y variados. Por ello, debe tenerse en cuenta que el **estado del arte** no se encuentra al mismo nivel y que la capacidad para garantizar ciertos requisitos quizás no puede ser la misma y deberá cubrirse en todo caso con garantías adicionales.

En este sentido, nos referimos a la ausencia de sesgo y a la explicabilidad de la Inteligencia Artificial.

En el ámbito científico, está asumido que ciertos problemas técnicos y sus soluciones no son posibles entenderlas y explicarlas correctamente. Por ejemplo, es común poner como ejemplo la prescripción de medicamentos: los médicos no comprenden realmente los efectos de muchas medicaciones (y tampoco quienes las consumen), pero las prescriben porque saben que ayuda a mejorar los síntomas de la mayoría de los pacientes pese a los efectos secundarios.

Algo similar ocurre con algunos modelos de Inteligencia Artificial. Existen por ejemplo sistemas basados en redes neuronales de gran complejidad técnica, lo cual complica su explicabilidad pero al mismo tiempo permite obtener resultados más precisos e informar de tasas de errores asociadas.

De lo anterior deben extraerse dos cuestiones:

1. Debe **garantizarse la calidad de los sistemas de IA, a través de la precisión y del control de las tasas de error**, donde se incluyen también aspectos como el sesgo racial, por género, etc.

Un modelo de IA que se encuentre en el estado del arte permite proporcionar a los usuarios del mismo unas garantías per se. Esa referencia al estado del arte debe hacerse no sólo respecto a tasas de falsos positivos y falsos positivos en precisión del sistema, sino también de no discriminación por cualquier sesgo.

Sin embargo, debe tenerse en cuenta que el estado del arte actual en algunos casos no permite garantizar un sesgo nulo o una precisión del 100%, pero sí muy reducido. Por ello,

sería necesario prever este detalle en la Carta de Derechos Digitales, la referencia a en la medida que sea posible de acuerdo con el estado del arte, para que no sea limitante para la aplicación de ciertas tecnologías.

Siguiendo esto, en las regulaciones específicas para los distintos casos de uso que se promulguen, sería recomendable, como ya ha comenzado a hacerse en España, que se especifiquen requisitos técnicos exigibles a los sistemas a emplear partiendo del nivel del estado del arte (p.ej. tasas de error máximas aceptables).

2. Debe **garantizarse la transparencia y no arbitrariedad de los sistemas de IA, si bien no son plenamente explicables.**

Como indicábamos, cuando se trata con algoritmos de IA hay ciertas cuestiones que pueden explicarse sobre su funcionamiento, pero otras no será posible. El usuario debe poder conocer que se va a tener en cuenta para la evaluación a realizar, pero quizás no sea posible explicarle cómo se va a realizar esa evaluación.

Por ello, resultaría también necesario hacer esa apreciación en la Carta de Derechos Digitales cuando se hace referencia a esta cuestión, y limitarlo a siempre que sea posible y de conformidad con el estado del arte. De otra forma, podría de alguna forma limitar el uso o buen uso de ciertas tecnologías más complejas (y muy precisas) en detrimento de otras quizás más explicables pero menos avanzadas.

Asimismo, a este respecto, podría plantearse la pregunta sobre la explicabilidad de las decisiones humanas en comparación con la de los sistemas de IA, lo cual es un tema del que se viene hablando mucho últimamente.

En definitiva, **el objetivo es que los modelos de Inteligencia Artificial no generen indefensión a las personas que los usan.**

Esto es algo que ya se plasmó en el Reglamento General de Protección de Datos, y que en esta Carta de Derechos Digitales queda perfectamente recogido en el apartado 2 del Título XXIII, amparando a las personas usuarias de los sistemas de Inteligencia Artificial a no ser objeto de decisiones basadas únicamente en procesos de decisión automatizadas si les van a producir efectos jurídicos o afectar significativamente. En estos casos, tienen a su disposición la posibilidad de impugnar la decisión del sistema automatizado y a solicitar la intervención humana.

De este modo, el sistema de IA podría asemejarse a la figura de un **perito**, que ayuda en el proceso aportando información, informes y evaluaciones en calidad de experto en la materia.

Al añadir las precisiones que se proponen en el presente documento, consideramos que no se reducen las garantías que se pretenden ofrecer a los ciudadanos en sus Derechos ante la Inteligencia Artificial, sino que las amplían al no limitar (al menos de forma interpretativa) la posible inclusión de ciertos sistemas avanzados y precisos que, por el estado del arte actual, quizás no puedan alcanzar algunos de esos requisitos, pero que en todo caso protegen al ciudadano porque siempre tendrá a su disposición la posibilidad de solicitar la intervención humana en el proceso y no ser objeto de decisiones únicamente automatizadas.

## Título XXIII - Derechos ante la Inteligencia Artificial

### Comentario

Según lo expuesto anteriormente, se considera apropiado incluir en el apartado 1 del Título XXIII una referencia al estado del arte, para que tengan cabida todos los sistemas de Inteligencia Artificial.

### Propuesta

Se propone modificar el apartado 1 del Título XXIII:

Donde dice:

“1. En el desarrollo y ciclo de vida de los sistemas de Inteligencia Artificial:

- a) Se deberá garantizar el derecho a la no discriminación algorítmica, cualquiera que fuera su origen, causa o naturaleza del sesgo, en relación con las decisiones y procesos basados en algoritmos.
- b) Se asegurarán la transparencia, auditabilidad, explicabilidad y trazabilidad.
- c) Deberán garantizarse la accesibilidad, usabilidad y fiabilidad.”

Se propone la siguiente redacción alternativa:

“1. En el desarrollo y ciclo de vida de los sistemas de Inteligencia Artificial:

- a) Se deberá garantizar, **en la medida de lo posible y de acuerdo al estado del arte**, el derecho a la no discriminación algorítmica, cualquiera que fuera su origen, causa o naturaleza del sesgo, en relación con las decisiones y procesos basados en algoritmos.
- b) Se asegurarán, **en la medida de lo posible y de acuerdo al estado del arte**, la transparencia, auditabilidad, explicabilidad y trazabilidad.
- c) Deberán garantizarse la accesibilidad, usabilidad y fiabilidad.”