

Carta de Derechos Digitales:

Propuesta del Consejo Asesor de Inteligencia Artificial

Miembros del CAIA que apoyan esta propuesta: Pilar Manchón, Lorena Jaume-Palasi, Emilia Gómez, Amparo Alonso Betanzos, Ricardo Baeza-Yates, Carme Torras

Introducción y Agradecimientos

La inteligencia artificial (IA) es una de las tecnologías de mayor proyección e impacto en la actividad humana en las próximas décadas. La IA actúa como catalizador de la innovación en el desarrollo tecnológico, con un fuerte impacto transformador en múltiples sectores de actividad.

Asimismo, La IA plantea importantes desafíos que se deben abordar de manera clara y meditada, y se han de establecer unos principios asociados al compromiso de desarrollo y uso de la tecnología de manera responsable.

Entendemos que el presente borrador es el fruto de un enorme esfuerzo por parte de sus autores, y esta propuesta alternativa ha sido elaborada con el respeto y el agradecimiento que dicho esfuerzo merece, y con el objetivo de enriquecer dicho trabajo con perspectivas expertas en distintos aspectos del campo de la IA.

Nuestro objetivo general puede sintetizarse de manera sencilla: ***“Proteger los derechos de la ciudadanía con respecto al uso de la IA, asegurándonos al mismo tiempo de que España pueda desarrollar sus capacidades en IA de manera competitiva en el mercado europeo y global.”***

Texto Alternativo Propuesto

1. Se velará por que el uso de sistemas de Inteligencia Artificial no magnifique los riesgos, prejuicios y perjuicios existentes en la sociedad.
2. En los contextos donde el uso de un sistema inteligente pueda dar lugar a un alto riesgo individual y/o colectivo, se definirán una serie de requisitos relacionados con su evaluación, transparencia y trazabilidad, supervisión humana y gobernanza, de acuerdo con las directrices propuestas por la Comisión Europea en su libro blanco de IA. **(1b, c, 2 y 3 en el original)**
3. La IA no debe violar los principios de buena fe o ser contraria al orden público o a las buenas costumbres. **(4 en el original)**

Comentarios y justificación de la propuesta

Observaciones técnicas sobre el formato de la Carta

Entendemos que la carta se ha formulado con un carácter ético, y es por ello que recomendamos usar “principios” y vocabulario formal de la semántica ética, y evitar explícitamente un formato y terminología legalistas. Por ejemplo, formalmente y debido a su carácter ético, integrar leyes ya existentes e incluirlas en la Carta como un principio ético significa, en la lógica y sistemática normativa tanto del derecho como de la ética, “rebajar” su estatus.

Artículo XXIII: Derechos ante la IA.

Aunque la carta consta de 24 artículos, esta propuesta se centra únicamente en el artículo XXIII. Nuestras propuestas se han basado en el siguiente **procédure**:

1. Interpretar la intención del artículo de la Carta relacionado explícitamente con la IA. ¿A qué problemas se refiere el artículo? ¿Qué quiere compensar o evitar?
2. Comprobar si ya existe ley que corresponda a esa intención.
3. Identificar qué lagunas normativas o incertidumbres dentro de la norma en concreto existen
4. Sopesar los efectos secundarios de la norma propuesta (¿es una regla demasiado ambigua que no funcionará en todos los contextos? ¿Crearán incentivos no deseados? ...)
5. Verificar que los artículos son adecuados desde el punto de vista técnico y de la investigación en IA.
6. Correspondencia con las iniciativas internacionales, por ejemplo con el libro blanco Europeo.
7. Equilibrio entre la confianza, la proporcionalidad de la norma y la excelencia en la IA

Punto 1. Creemos que el impacto de los algoritmos, a parte de los riesgos relacionados con la discriminación, debe ampliarse a otros prejuicios y riesgos. También creemos que debe indicar que el uso de la IA no los magnifique, porque pueden ya existir por otras causas (nuevo punto 1).

La IA puede afectar a otros derechos fundamentales como derechos del trabajador, de libertad de expresión, incluso a la integridad física de la persona (si se tratan de sistemas utilizados, por ejemplo, en la salud). Proponemos que en el punto 1 se formulen principios que cubren sectores o riesgos para los que no hay leyes de protección o que requieren adaptación debido a la IA. P. ej.

1. Hay sectores o riesgos que la ley aún no cubre (por ejemplo la ley de igualdad no se aplica a todos los sectores, como el derecho financiero),
2. Hay problemas estructurales/sociales (como el impacto medioambiental, la cohesión social, etc.) que no están cubiertos por los derechos fundamentales y que a nivel legal tampoco están debidamente desarrollados.

Punto 2. (puntos 1 b y c, 2 y 3 en el original). En nuestra opinión, los requisitos de transparencia y supervisión humana se pueden resumir dentro del concepto de gobernanza (puntos 1,2,3). Creemos que, además de transparencia, se necesitan otras medidas como un análisis de impacto, auditorías, procesos de acreditación, etc. Es decir, proponemos reflejar un repertorio más amplio de medidas de evaluación y supervisión para los que la transparencia y trazabilidad dependiendo del contexto, tienen un propósito determinado que no es siempre el mismo. En según qué casos, la transparencia sirve para informar al consumidor; en otros casos, es necesaria para poder efectuar una auditoría por parte de una

entidad determinada; adicionalmente, se puede requerir transparencia no sobre el sistema, sino sobre los mecanismos para poder interpelar o exigir asistencia. Eso debe especificarse para evitar que la transparencia se convierta en una medida que transfiere la responsabilidad al usuario en contextos en los que esa transferencia de responsabilidad no es legítima. Asimismo, es importante contextualizar los procesos y reglas de esa gobernanza y evitar reglas al estilo one-size fits-all que aplican las mismas reglas para PYMES y grandes compañías - o para sistemas en servicios de un hospital y sistemas que se usan para controlar las persianas de una casa. Por eso especificamos y recomendamos concentrarse en sistemas “de alto riesgo”. Queda por definir criterios procedimentales con los que determinar cuándo un sistema es de alto riesgo. La inclusión del requisito de explicabilidad creemos que es difícil de implementar, ya que muchos de los algoritmos de IA actuales carecen de ese requisito. La transparencia y trazabilidad creemos que refleja de forma más precisa las necesidades al respecto.

Punto 3 (punto 4 original). Interpretamos el punto 4 como una línea roja. Para las líneas rojas que son difíciles de sistematizar, el lenguaje normativo usa tradicionalmente conceptos como principios de buena fe, el orden público o las buenas costumbres. Estos principios son utilizados en casos en los que es difícil ofrecer una definición y sistematización que pueda ser aplicada de forma general. Con nuestra propuesta proponemos recurrir para este tipo de casos al poder judicial, como instancia normativa que permite mayor contextualización.