

## UNA IA ÉTICA: la UE y la gobernanza algorítmica

**Carme Colomina**, investigadora sénior, CIDOB

*La Unión Europea ha aprobado la primera ley de inteligencia artificial del mundo con reglas vinculantes que impondrán obligaciones y límites a los distintos usos de la IA según el riesgo que comporten. La inteligencia artificial, y su desarrollo hiperacelerado, suponen un desafío tecnológico, económico, social y geopolítico, que los grandes poderes globales enfrentan desde visiones estratégicas a veces contradictorias. A estas distintas aproximaciones se suman las múltiples evidencias de las desigualdades originadas o reforzadas por el uso de modelos y programas de aprendizaje automático, que obligan a despejar incertidumbres y garantizar la rendición de cuentas.*



# 784

DICIEMBRE  
2023

CIDOB **50**  
years  
BARCELONA  
CENTRE FOR  
INTERNATIONAL  
AFFAIRS

 Cofinanciado por  
la Unión Europea

La ubicuidad de la Inteligencia Artificial es indiscutible. Su capacidad de penetración abarca desde la seguridad del ciberespacio a la robótica doméstica; desde la imitación de la inteligencia humana al poder de decisión sobre préstamos bancarios o ayudas sociales. La IA parece infiltrarlo todo. Pero su desarrollo tecnológico avanza a un ritmo tan vertiginoso que no permite despejar incertidumbres. La gobernanza de la IA plantea problemas conceptuales, funcionales, analíticos, de orden práctico y jurídico. Sin embargo, arrastrados por esta temporalidad hiperacelerada, los grandes actores globales, públicos y privados, se han puesto a discutir sobre cómo regular la IA cuando todavía estamos intentando aprehender bien qué es y cuál es su verdadero alcance. Por eso fallan los consensos.

Sin embargo, la Unión Europea ha tomado la iniciativa y ha conseguido cerrar un acuerdo interinstitucional, a finales de 2023, para aprobar la primera Ley de Inteligencia Artificial que pretende limitar los riesgos en el uso de la IA estableciendo distintas categorías (inaceptables, elevados, limitados o mínimos). Se consideran inaceptables y, por tanto, se prohíbe el uso de sistemas de reconocimiento de emociones; los sistemas de manipulación intencionada para influir en el comportamiento de las personas y causar daños físicos; y los sistemas que explotan vulnerabilidades de un grupo específico de personas. Otro de los temas espinosos de la negociación fue la inclusión de la IA generativa (los modelos fundacionales), que deberán cumplir con requisitos de transparencia. Todos los actores y usuarios de la IA en la Unión Europea deberán respetar un Código de Conducta ético respaldado por multas, ya que la ley prevé un sistema de sanciones, que puede llegar a un porcentaje del volumen total de negocios de la compañía infractora.

Ha sido una negociación larga, que deberá ratificarse en 2024 y contará con dos años de margen para su implantación gradual. La UE se convierte así en pionera en la definición de un marco legal. Sin embargo, también hay voces críticas que consideran que el esfuerzo se queda corto ya que se centra en el riesgo de los posibles usos de la IA, pero no soluciona los daños inherentes a la propia tecnología. Es imprescindible complementar los enfoques centrados en encontrar «soluciones a los problemas» –basados en minimizar los daños causados por la IA– con **visiones centradas en el «diagnóstico de problemas»**, es decir que examinen el origen y las causas de los actuales imaginarios, y de la economía basada en la IA. Como se pregunta Urvashi Aneja, directora de Digital Futures Lab, **«¿es suficiente un enfoque meramente regulatorio, cuando lo que está en juego es cómo conocemos y entendemos el mundo?»**. La Unión Europea asegura, por su parte, que el reglamento «tiene como objetivo garantizar que los sistemas de IA utilizados en la UE sean seguros y respeten los derechos fundamentales y los valores europeos».

*La UE se convierte así en pionera en la definición de un marco legal. Sin embargo, también hay voces críticas que consideran que el esfuerzo se queda corto ya que se centra en el riesgo de los posibles usos de la IA, pero no soluciona los daños inherentes a la propia tecnología.*

De hecho, la UE y sus estados miembros se encuentran en **una situación de difícil equilibrio** como reguladores y, a la vez, usuarios de los sistemas automatizados de toma de decisiones en el sector público. Por un lado, los gobiernos europeos deben proteger a los ciudadanos de los posibles daños de las tecnologías emergentes, pero también, y cada vez más, utilizan estas tecnologías para gobernar, en un contexto de plena consciencia sobre **los efectos devastadores** que los sesgos algorítmicos pueden tener sobre las predicciones, recomendaciones y decisiones públicas automatizadas. Es imposible, a estas alturas, negar el impacto discriminatorio de muchos de estos sistemas, aunque los gobiernos, en general, se nieguen a **la transparencia que reclaman las organizaciones civiles**.

**Como defiende Lorena Jaume-Palasi**, los sistemas algorítmicos centrados en perfiles humanos crean estándares, pero un perfil humano es ya en sí una distorsión del individuo. Los datos no son neutrales y solo son parcialmente representativos. En este contexto, la inteligencia artificial y los sistemas de gobernanza algorítmica han acabado reproduciendo sesgos discriminatorios. De hecho, los sesgos son sociales y, si no se corrigen, la regulación puede acabar teniendo un efecto legitimador y amplificador.

## Los riesgos del debate normativo

En este debate sobre la gobernanza de la IA, que va de la mano de la carrera geopolítica por liderar la innovación tecnológica, la Unión Europea, una vez más, lo fía todo a su credibilidad reguladora. 2024 será un año fundamental en todo el debate normativo, a pesar de que la falta de consensos entre grandes actores llega a un nivel tan básico como la propia definición de la Inteligencia Artificial. Algunos reguladores han optado por definiciones amplias, por temor a que el texto quedara obsoleto a corto plazo, como en los «Principios de la OCDE sobre Inteligencia Artificial». Mientras, otros, en cambio, han preferido referirse a las características de la IA sin definirla al detalle.

*2024 será un año fundamental en todo el debate normativo, a pesar de que la falta de consensos entre grandes actores llega a un nivel tan básico como la propia definición de la Inteligencia Artificial.*

Pero no se trata solo de si debe regular o no, sino de quién debe hacerlo. Se vio perfectamente en la [primera cumbre mundial sobre seguridad de la IA](#), convocada el 1 de noviembre por el primer ministro británico, Rishi Sunak. Un punto de encuentro de los grandes poderes globales –públicos y privados; tecnoautoritarios o abiertos– intentando regular o influir en los debates sobre regularización en curso. La «Declaración Bletchley», firmada por 28 países, recoge un compromiso para abordar las principales amenazas de la IA, un acuerdo para examinar los modelos de IA de empresas tecnológicas antes de su lanzamiento, así como un pacto para establecer un panel global de expertos sobre inteligencia artificial. Además, en la embajada de Estados Unidos en Londres, 31 países firmaron, en paralelo, una declaración (no vinculante) para establecer límites al uso militar de la Inteligencia Artificial.

Pero esta cumbre global también ejemplifica uno de los riesgos del debate normativo: la potencia de las grandes empresas tecnológicas dominantes –y su esfuerzo por liderar el debate sobre los límites de los marcos reguladores de la misma manera que marcan el ritmo de la aceleración tecnológica– por encima de las voces de la sociedad civil. Es una confrontación no exenta de una nueva forma de colonialismo; de explotación de los usuarios, convertidos en una fuente infinita de datos a cruzar, analizar y comercializar. La protección y fiabilidad de los datos, por ejemplo, debería incluirse desde el mismo diseño de la tecnología. Y no sólo eso, debería tenerse en cuenta desde el mismo momento de la programación quien está siendo excluido de la toma de decisiones o de los datos que se utilizarán para la automatización de la toma de decisiones.