

# Inteligencia artificial y derechos humanos: ¿se necesitan mutuamente?

Karlos A. Castilla Juárez



Agentes de la policía de Bosnia, en la estación de tren de Bihać, devuelven a un grupo de migrantes hacia Sarajevo en autobús, el 10 de noviembre de 2018. Fotografía: [Alessio Mamo](#)

Al hablar de la relación entre derechos humanos e inteligencia artificial (IA), lo más común actualmente es preguntarse cómo la IA afecta a los derechos humanos. De hecho, existe un grupo concreto de derechos humanos que constantemente se señala como el más expuesto a estas afectaciones; por tanto, su protección en la gobernanza y la regulación de la inteligencia artificial se considera prioritaria. Es el caso de la privacidad y la protección de datos, la libertad de expresión y el acceso a la información, la igualdad y la no discriminación.

Que estos derechos humanos y no otros estén constantemente en el centro de atención parece normal si se entiende que, para su diseño, desarrollo y despliegue, la inteligencia artificial necesita datos, que son la materia prima con la que alimenta los algoritmos para poner en marcha cualquier IA. Es por ello que conocer la forma en que se obtienen estos datos (si es con o sin autorización), cómo se almacenan y cómo se utilizan, qué parte se utiliza o se excluye, cuáles no se recogen, cómo se difunden o comparten, etc., son aspectos de gran importancia. Al final, los datos son información relevante sobre la vida y las

actividades que desarrollamos las personas.

Ante esto, lo que menos nos hemos preguntado es si los derechos humanos necesitan la inteligencia artificial. Es decir, si es necesario disponer de la IA para ejercer algunos o todos nuestros derechos humanos, porque es aquí donde el número de derechos humanos susceptibles de estar afectados por la inteligencia artificial se ampliaría, puesto que no sólo estaríamos hablando temas en torno a los datos, sino que se podría incidir en el disfrute, ejercicio y garantía de otros derechos.

Así, por ejemplo, ¿necesitamos la inteligencia artificial para garantizar el derecho a la vivienda? ¿Necesitamos la IA para ejercer la libertad de circulación? ¿Necesitamos la IA para ejercer y garantizar el acceso a la justicia? La respuesta a estas preguntas, si se piensa en la pretendida universalidad de los derechos humanos, no es sencilla y difícilmente puede ser absoluta en sentido positivo o negativo, por lo que es importante reflexionar en torno a esto.

## ¿Cuál es la relación entre derechos humanos e inteligencia artificial?

En 2023 bien podemos afirmar que la relación entre derechos humanos e inteligencia artificial es más de una “vecina lejana” que de un “compromiso mutuo”. Esto es así porque aunque derechos humanos e IA viven en el mismo mundo y se conocen en términos genéricos, han interactuado poco de forma cercana, a pesar de los múltiples esfuerzos, especialmente desde organismos internacionales de derechos humanos, [1] para que se conozcan mejor y colaboren mutuamente.

Quienes trabajamos en el ámbito de los derechos humanos, con el paso de los años —y al ver cómo la IA llegaba a más ámbitos del ejercicio, disfrute y garantía de estos derechos— nos hemos ido acercando, muchas veces sin entender en detalle y en profundidad todo lo que implica. Por otra parte, quienes trabajan en el sector de la inteligencia artificial se han interesado poco en los derechos humanos y han preferido mirar a la ética cuando se les han cuestionado los efectos negativos que puede tener la IA en la vida de las personas.

Por todo ello, podemos afirmar que la relación entre la IA y los derechos humanos todavía está en construcción, a pesar de su gran relevancia: prácticamente en todos los ámbitos donde se está implementando la inteligencia artificial existe un derecho humano vinculado, directa o indirectamente. Sí, esa afirmación no es exagerada. El problema es que muchas veces no vemos esta vinculación porque perdemos de vista que, por ejemplo, al pretender utilizar un transporte público donde existen aplicaciones de IA se podría estar incidiendo en el ejercicio y garantía de nuestros derechos humanos a la libertad de circulación, privacidad, no discriminación o reconocimiento de la personalidad jurídica, entre otros. Y así en muchas actividades cotidianas.

| La relación entre la inteligencia artificial y los derechos humanos

todavía está en construcción, a pesar de su gran relevancia: prácticamente en todos los ámbitos donde se está implementando la inteligencia artificial existe un derecho humano vinculado

En este sentido, la relación entre derechos humanos y la inteligencia artificial es muy importante y debe ser prioritaria. Es necesario que haya más interacción entre la comunidad de derechos humanos y de la IA para que el desarrollo tecnológico no se cree sin la comunidad de derechos humanos. Especialmente porque la ética a la que se apela como forma de autorregulación desde la inteligencia artificial (empresas, centros de investigación, academias de ciencias, así como algunos estados, organizaciones internacionales y asociaciones de la sociedad civil) es insuficiente para prevenir y evitar los posibles efectos negativos de la IA en las personas.

Así, resulta evidente que en la relación entre la IA y los derechos humanos queda mucho por hacer y que, en el futuro próximo, debería ser una relación de compromiso mutuo por todo lo que se explicará a continuación.

## Derechos humanos en el ámbito de la inteligencia artificial: ¿dónde queda la ética?

En el mundo, quien más IA produce son las empresas, entre las que destacan Google, Amazon, Apple, Microsoft, Meta, Baidu, Alibaba y Tencent. Esto significa que el grueso de la inteligencia artificial se está produciendo en Estados Unidos y China, o bajo intereses económicos de personas con estas nacionalidades. De ahí que parezca evidente que son estas empresas las que han ido estableciendo en la práctica su propia regulación y el marco de actuación vinculado con la IA. Esto, sin duda, influye también en pequeñas empresas y en toda la generación (o mercado) de la IA. Por este motivo, justamente desde muchos ámbitos vinculados con sectores dominantes de la IA (empresas, centros de investigación, academias especializadas y asociaciones vinculadas con todos los ámbitos anteriores) se propone insistente la ética como marco óptimo de regulación que puede garantizar derechos y no pone en riesgo la libertad de empresa, la propiedad intelectual y, por lo general, sus legítimos intereses económicos, de innovación o de negocio.

Si no fuera así, no se entendería por qué Europa —donde el poder económico y político de quienes producen IA es aún menor— es la única región del mundo donde, en 2023, se está avanzando en una regulación de la IA con una mirada más cercana a los derechos humanos, abandonando las simples regulaciones éticas. Es cierto que existen obstáculos, y que ciertos aspectos requerirán análisis más completos y ajustes, especialmente si uno de los objetivos es la protección de los derechos fundamentales de las personas.

En este contexto, la ética en la inteligencia artificial que se impulsa y quiere establecerse como marco de regulación y prevención de afectaciones a las personas se muestra como cambiante, maleable, parcial e incierta, y permite la impunidad en casos de

incumplimiento. Esto es especialmente así cuando son las grandes empresas, con sus intereses económicos o de negocio, las que finalmente establecen esta ética o las que proponen sus parámetros y ámbitos de aplicación.

Ante esto, tenemos los derechos humanos, que, en primer lugar, nos permiten presuponer que se han resuelto cierto conjunto de debates filosóficos para concretarlos e identificarlos. En segundo lugar, representan acuerdos mínimos sobre valores universales (aunque no lo sean todos) en forma de derechos y libertades, por lo que sabemos con mucha certeza qué derechos existen, [2] qué alcance tienen, dónde están reconocidos y quién debe respetarlos, protegerlos y garantizarlos.

Además, el apoyo mundial a estos derechos es bastante amplio: existen aspectos comunes en normas nacionales e internacionales de muchas regiones del mundo. Así, estas empresas (y otras que también producen IA), como cualquier otra persona, tienen responsabilidades frente a los derechos humanos de los demás a partir del contenido que establecen las leyes a nivel nacional en materia civil, laboral, administrativa y penal, y también con las derivadas de normas nacionales e internacionales de derechos humanos, aunque esto se niegue constantemente y se eluda por parte de todo tipo de empresas; tanto es así que hubo la necesidad de desarrollar los principios rectores sobre las empresas y los derechos humanos de Naciones Unidas como un intento de reafirmar este conjunto de mínimos a observar.

Así pues, los derechos humanos en el ámbito de la inteligencia artificial deben tener, en primer lugar, una incidencia principal en todo lo relacionado con los datos que utiliza la IA para los derechos que podrían verse afectados (la privacidad y la protección de datos, la libertad de expresión y el acceso a la información, la igualdad y la no discriminación). En segundo lugar, también deben tener influencia en los ámbitos en los que puede ser aplicada la IA, es decir, respecto a las preguntas y respuestas que los diferentes algoritmos darán como resultado, ya que es aquí donde se puede incidir en muchos más derechos.

Si nos fijamos en la primera perspectiva vinculada a los datos, por ejemplo, siempre será mejor aplicar los estándares internacionales de derechos humanos que se han ido creando, desarrollando y actualizando respecto al derecho a la vida privada, ya que tienen unos parámetros claros cuando se trata de personas públicas y personas comunes, y tienen definida qué parte de este derecho no puede transgredirse en ningún caso. Aplicar estos estándares es preferible a dejar que una empresa o persona valore o evalúe éticamente qué es lo que se considera “éticamente adecuado” que se sepa, divulga o utilice de la vida privada de una persona o de un conjunto de personas sobre las que se han recogido una serie de datos con su o sin consentimiento total.

Desde la segunda perspectiva, y como ya se intenta hacer en la propuesta de regulación europea, es necesario definir los ámbitos en los que los derechos humanos se encuentran en riesgo evidente y, por tanto, también los parámetros sobre los que se debe limitar el uso y desarrollo de la inteligencia artificial. En el caso de Europa, por ejemplo, a partir del establecimiento de “niveles de riesgo”, [3] todos los sistemas de IA que supongan una amenaza clara para la seguridad, los medios de vida y los derechos de las personas se

consideran un riesgo inaceptable: desde la puntuación social por parte de los gobiernos hasta los juguetes que hacen uso de asistentes de voz, las cuales fomentan comportamientos peligrosos. Hacer valoraciones éticas *ad hoc* de estas cuestiones no garantiza unos mínimos de certeza y seguridad jurídica, aunque estas valoraciones se hagan con buena voluntad.

Así, en la Unión Europea, a partir del establecimiento de ámbitos de alto riesgo, riesgo limitado y riesgo nulo, se ha propuesto que se prohíba el uso de la inteligencia artificial en sistemas de categorización biométrica que realicen servir a características personales sensibles como el género, la raza, la etnia, la religión o la orientación política. Asimismo, estarían prohibidos los sistemas predictivos de vigilancia para evaluar el riesgo de una persona o grupo de personas de cometer un delito u ofensa, así como los sistemas de reconocimiento de emociones por parte tanto de la policía como de los agentes de fronteras, o en puestos de trabajo y escuelas. De igual forma, se limitará el rastreo aleatorio de datos biométricos de redes sociales o cámaras de vigilancia para crear o ampliar bases de datos de reconocimiento facial. Con todo, parece claro que existen ámbitos de los derechos humanos donde claramente la IA debe estar vetada, por más que éticamente alguien pudiera considerar admisible que entrara.

La ética como forma de autorregulación de la inteligencia artificial es insuficiente para prevenir los efectos negativos en las personas

Desde una perspectiva empresarial, de negocio, estrictamente económica y mercantilista capitalista, el establecimiento de ámbitos prohibidos para la IA puede considerarse una intromisión grave. De ahí, en parte, la insistencia por apostar por las regulaciones éticas. Pero desde la perspectiva de la protección de los derechos humanos, es evidente que hay unos mínimos de estos derechos y libertades que deben ser innegociables, al igual que también existen límites para no vulnerar los derechos humanos en muchos otros ámbitos en el “mundo analógico” en el que hemos vivido hasta ahora.

## La inteligencia artificial en el ámbito de los derechos humanos

En el sector de la inteligencia artificial, los espacios en los que hay que tener en cuenta los derechos humanos parecen claros. Ahora, sin embargo, es necesario analizar una perspectiva menos explorada: en el ámbito de los derechos humanos, ¿es necesario que la IA esté presente? O, dicho de otro modo, ¿es necesaria la IA para el ejercicio de algunos o de todos nuestros derechos humanos?

Si tomamos en consideración el hecho de que, en su origen, los derechos humanos no necesitaron la inteligencia artificial para su ejercicio, garantía y protección, en 2023, y de cara al futuro, no parece imprescindible el uso de esta tecnología para estos fines, salvo

que la inteligencia artificial permita asegurar la pretendida universalidad de los derechos humanos que, 75 años después de la aprobación de la Declaración Universal de los Derechos Humanos, sigue siendo un tema pendiente en todas las regiones del mundo.

En cualquier caso, determinar si la inteligencia artificial es necesaria para el ejercicio, garantía y protección de los derechos humanos no puede partir de una conclusión general para todos los derechos, sino que requiere el análisis de cada uno de los derechos reconocidos a nivel nacional e internacional para determinar, a partir de sus elementos y caracterización, si la IA es necesaria o útil. Incluso requiere un análisis más particular del lugar o país en cuestión, ya que existen lugares en el mundo donde el acceso a internet no está garantizado, con lo que pensar en la IA sería multiplicar las desigualdades. [4]

El análisis particular de los derechos es muy importante porque, por ejemplo, no es lo mismo el ejercicio del derecho a la protesta pacífica que el ejercicio del derecho a la vivienda. Protestar podríamos hacerlo con impactos idénticos a los del mundo analógico o, incluso, podría pensarse, en algunos casos, en efectos más focalizados, si se hace en entornos virtuales o digitales, mientras se cumpla con las finalidades del derecho. Pero siempre sería insuficiente disponer de una vivienda en estos mismos entornos si en la realidad analógica no se tiene o si ni siquiera se puede acceder porque no podemos utilizar la aplicación que nos lo permitiría, o porque la aplicación citada nos excluye sin ninguna justificación razonable por las características y datos con los que funciona el algoritmo.

En este sentido, es relevante señalar que el hecho de que no exista una inteligencia artificial determinada para el ejercicio, la garantía y la protección de un derecho específico no debe considerarse *per se* como un aspecto negativo. Debemos asumir y entender que no todos los derechos humanos necesitan el uso de la IA para ser efectivamente ejercidos, garantizados y protegidos. Más bien lo que ocurre es que si en el mundo analógico hay personas que están siendo excluidas o no pueden ejercer un derecho determinado, con mucha probabilidad también será así con el uso de aplicaciones de inteligencia artificial.

Fijémonos, por ejemplo, en el acceso a la justicia. En el mundo real analógico, actualmente muchas personas no tienen garantizado de forma efectiva este derecho por múltiples razones, incluso en países llamados desarrollados, [5] por lo que la implementación de una «justicia digital» o de «tribunales automatizados» no es en sí misma una solución a ese problema. Menos aún si el acceso a estas aplicaciones digitales debe hacer frente a las mismas barreras de acceso que la justicia analógica. [6] Por tanto, muy posiblemente tan sólo es una buena opción para quienes ya accedían a la “justicia analógica”.

Sin embargo, tampoco supone necesariamente una solución para quienes ya accedían a la justicia si esta “justicia digital” es incapaz de asegurar las garantías judiciales que se requieren en todo proceso (la imparcialidad del juez, la publicidad del proceso, la posibilidad de asistencia de un abogado, la prohibición de dilaciones indebidas y la utilización de los medios de prueba disponibles, entre otros). Así pues, debemos tener cuidado con lo que muchas veces se promueve como “modernidad”, porque no necesariamente implica acceder a los derechos humanos, ejercerlos y respetarlos de manera efectiva.

Es importante decidir el uso de la inteligencia artificial pensando en los derechos humanos. Si sólo pensamos en la modernidad, existen muchas probabilidades de que simplemente se estén creando formas modernas de exclusión

Ante esta situación, es importante decidir el uso de la inteligencia artificial pensando en los derechos humanos, para valorar si la IA aporta o no algo mejor a lo que ya existe en el mundo analógico para su ejercicio, garantía y protección. Es decir, no hay que tener miedo a decir que no a una aparente modernidad que no aportará nada positivo a los derechos humanos, y mucho menos cuando quien lo ofrece son empresas a las que importa su negocio y no la universalidad de los derechos humanos.

En el ejemplo anterior, puede haber aspectos concretos de este acceso a la justicia o tutela judicial en los que la IA resulte útil, como podría ser, por ejemplo, la gestión de expedientes, el acceso a estos expedientes, la gestión de información o la solución de problemas simples. [7] Pero esto no significa que sea útil para todo lo que implica y debe contener el derecho.

Tampoco podemos esperar que toda decisión respecto a un derecho humano recaiga en la inteligencia artificial, ya que, por más datos que haya, la intervención humana siempre será relevante en la asignación individual de derechos y libertades. Por todo ello, parece innegable que la IA puede ayudar con consejos, sugerencias y parámetros a la hora de tener una mejor comprensión de situaciones o conseguir información relevante, pero la decisión final no debe depender exclusivamente de esta tecnología.

Para determinar si necesitamos la IA para el ejercicio, garantía y protección de un derecho específico o de una libertad concreta, es recomendable plantearse, al menos, las siguientes preguntas:

- Qué aporta la IA al *derecho X* ¿que vaya más allá de su ejercicio, garantía o protección en la realidad analógica?
- ¿Qué nivel de acceso efectivo tiene en estos momentos el *derecho X* en la realidad analógica?
- El uso de la IA mejora el acceso o ejercicio del *derecho X* y no perpetúa las exclusiones existentes (o, incluso, crea nuevas)?
- ¿La IA es capaz de garantizar todos los elementos y caracterizaciones que se le han dado al *derecho X* en los estándares de derechos humanos?
- ¿Cuál es el nivel de decisión que se delegará a la inteligencia artificial para el ejercicio, garantía o protección del *derecho X* y, por tanto, cuál es el nivel de decisión que tendrán las personas?

Si antes de decidir que el derecho a la salud, el derecho al trabajo, el derecho a la educación, la libertad de expresión, el derecho de acceso a la información, el derecho a la nacionalidad o la libertad de circulación se ejercerán, se garantizarán o se protegerán por medio de la inteligencia artificial nos planteamos las preguntas anteriores, posiblemente se podrán reducir o evitar nuevas afectaciones a los derechos humanos. De lo contrario, si sólo pensamos en la «modernidad», existen muchas probabilidades de que simplemente se estén creando formas «modernas» de exclusión y vulneración de derechos humanos.

## Conclusión

Respecto a los derechos humanos, lo que nos puede preocupar de cualquier tecnología, no sólo de la inteligencia artificial, es, por un lado, el proceso de creación y diseño de tecnologías y, por otro, los usos, aplicaciones, creaciones, efectos o resultados de las tecnologías que excluyan, restrinjan, discriminan, anulen, menoscaben, impidan o reduzcan el reconocimiento, disfrute o ejercicio, en condiciones de igualdad, de los derechos humanos y las libertades fundamentales. [8]

Desde esta perspectiva, es importante no perder de vista los derechos humanos en la inteligencia artificial, en la misma medida que la IA en el ámbito de los derechos humanos, puesto que en ambas direcciones se producen incidencias, aunque no todas sean siempre negativas.

Si el objetivo es salvaguardar los derechos y libertades de las personas, la realidad actual nos muestra que los derechos humanos son más necesarios en el ámbito de la inteligencia artificial que la IA en el ámbito de los derechos humanos.

## REFERENCIAS Y NOTAS

- 1 — Véase, por ejemplo: Comisaria de Derechos Humanos del Consejo de Europa (2023). *Human rights by design future-proofing human rights protection in the era of AI*. Estrasburgo: Consejo de Europa.
- 2 — Risse, Mathias (2018). “Human Rights and Artificial Intelligence: An Urgently Needed Agenda”. *HKS Faculty Research Working Paper Series*. RWP18-015, p. 10.
- 3 — Comisión Europea (2020). *White Paper on Artificial Intelligence: a European approach to excellence and trust*. Bruselas.
- 4 — Castilla, Karlos (2022). *Cuatro ángulos de análisis de la igualdad y la no discriminación en la inteligencia artificial*, Barcelona: Instituto de Derechos Humanos de Cataluña, p. 6.
- 5 — Deseau, Arnaud; Levai, Adam; Schmiegelow, Michèle (2019). “Access to Justice and Economic Development: Evidence from an International Panel Dataset”. *LIDAM Discussion Papers IRES*, 2019009. Lovaina: Universidad Católica de Lovaina, Institut de Recherches Economiques et Sociales (IRES).
- 6 — Castilla, Karlos (2012). *Acceso efectivo a la justicia: elementos y caracterización*. México: Porrua

- 7 — Reiling, Dory (2020). "Courts and Artificial Intelligence". *International Journal for Court Administration* 8. 11(2), p. 10.
- 8 — Pont, Anna; Passera, Agostina; Castilla, Karlos (2022). "Análisis introductorio". *Impactos de las nuevas tecnologías en los derechos humanos*. Barcelona: Instituto de Derechos Humanos de Cataluña, p. 11.

**Karlos A. Castilla Juárez**

Karlos A. Castilla Juárez es abogado por la Universidad Nacional Autónoma de México y doctor en Derecho por la Universidad Pompeu Fabra. Es responsable de investigación del Instituto de Derechos Humanos de Cataluña, miembro del Observatorio de Derecho Público de la Universidad de Barcelona y profesor asociado de Derecho Internacional de los Derechos Humanos en la Universidad Pompeu Fabra. Su actividad investigadora se ha desarrollado en el ámbito del derecho internacional (sistemas internacionales de derechos humanos y litigio internacional), del derecho constitucional (control de convencionalidad y sistemas de acceso a la justicia), derechos de las personas migrantes (detención migratoria, expulsión, migración y cambio climático) y temáticas específicas de derechos humanos (racismo, igualdad, no discriminación, interseccionalidad, libertad de expresión y justicia transicional).