

IA Bajo el Volcán

MANUAL DE DEFENSA ALGORÍTMICA

Breve guía práctica para
conocer, entender y ejercer
tus derechos ante la IA

Elaborado por Esther Paniagua
para Prodigioso Volcán



Prodigioso Volcán

06/2025

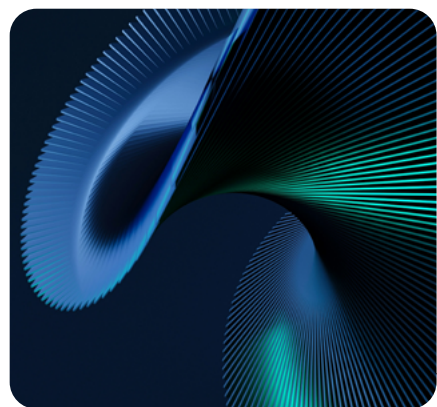
ÍNDICE

03

Introducción

08

El derecho a saber, entender
y actuar



17

El poder de la defensa
algorítmica: batallas ganadas

26

¿Qué exige la ley?

29

Las empresas:
aliadas clave

32

Hacia una defensa
algorítmica más activa

34

Referencias

INTRODUCCIÓN

¿Te has preguntado por qué no te eligieron para ese puesto? ¿O por qué te denegaron una hipoteca? ¿Por qué no te concedieron aquella ayuda pública? En muchos casos, detrás de estas decisiones puede haber un sistema automatizado evaluándote sin que lo sepas.

Un modelo predictivo basado en IA pudo haber arrojado un resultado que determinó que no te seleccionasen. ¿Tenía razón? Tal vez sí. Pero también puede que no, y nunca podrás saberlo.



En esta guía práctica encontrarás claves claras sobre

01

Cómo te afecta la IA
(aunque no la veas).

02

Qué derechos tienes
frente a los sistemas
automatizados.

03

Qué hacer para actuar en
tu defensa y reclamar.

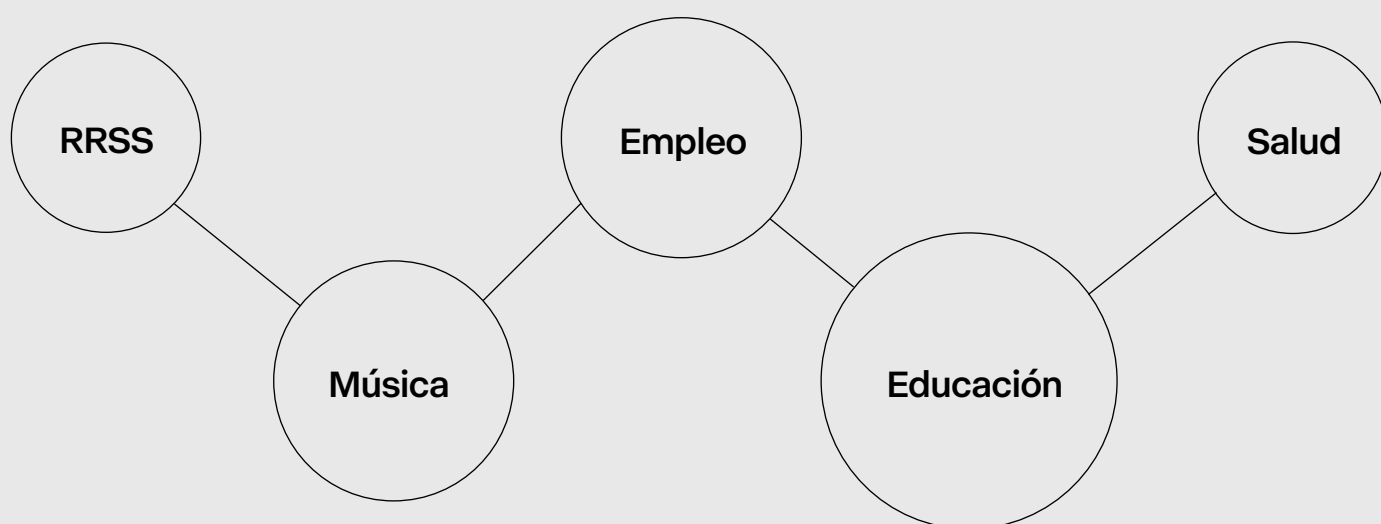
En España, los algoritmos ya deciden -o más bien se usan para decidir- quién recibe una ayuda de vivienda o una hipoteca, o el bono social; quién tiene prioridad para una prueba médica; qué víctima de violencia de género recibe protección; qué estudiante logra entrar al grado y la universidad deseadas, o quién comete fraude.



¿Por qué es importante?

A menudo, el uso de este tipo de sistemas es invisible, y pasa desapercibido. Por eso, reivindicar, ejercer y defender nuestros derechos ante la inteligencia artificial es tan importante. Porque lo que no se ve, no se conoce y no se entiende, no se puede cuestionar. Hablamos del derecho a conocer y entender la IA, y a reclamar y actuar si sentimos que se ha cometido una injusticia o se han violado nuestros derechos.

Este **Manual de defensa algorítmica** es, al mismo tiempo, una llamada a la acción y una guía para reclamar y proteger tus derechos frente a los algoritmos que ya influyen —y van a seguir influyendo— en tu vida.



Cada vez más aspectos de nuestra vida están mediados por sistemas automatizados, desde los más cotidianos hasta los más decisivos.

Contenido, entretenimiento, series, música, ficción, prensa... pero también empleo, créditos, seguros, redes sociales, salud, educación, ayudas públicas, policía predictiva o control migratorio. Son solo algunos de los ámbitos donde se usan algoritmos y modelos de IA para tomar decisiones que pueden condicionar, limitar o incluso impedir el acceso a oportunidades vitales.

Convertirnos en una **ciudadanía informada en IA** consiste en tres puntos básicos:

01

En **saber** cuándo y de qué forma se están utilizando algoritmos o sistemas automatizados en decisiones que nos afectan;

02

en **entender** cómo funcionan y cómo llegan estos sistemas a cada resultado que devuelven;

03

y en tener el poder de **actuar** cuando algo no encaja, sospechamos que puede haber un error o no estamos de acuerdo con la decisión tomada.

Estos tres pilares (saber, entender y actuar) son la base de una ciudadanía consciente y autónoma, con capacidad de decisión y con herramientas para defenderse en la era de la IA.

SABER ENTENDER Y ACTUAR

Cómo te afecta la IA

En tus interacciones con el móvil, la tablet y el ordenador y en muchas de las que llevas a cabo en el mundo físico. En tu vida digital y en tu vida analógica.



La IA interviene en decisiones sobre el contenido que ves en redes o plataformas, la publicidad que recibes, los precios que se te muestran al comprar online, las rutas que sigues, las ofertas laborales que te llegan (o no); las decisiones que toman bancos, aseguradoras o administraciones sobre ti, y un largo etcétera.



Para colmo, muchas veces, estas decisiones están basadas en datos que has dejado en alguna app o web, o en **el rastro que dejas online sin ser consciente de ello**, procesados por algoritmos invisibles en base a criterios incomprensibles. La IA está muy presente en el entorno que te rodea y, te des cuenta o no, influye en oportunidades, derechos y decisiones que pueden marcar tu vida.

¿Qué podemos hacer desde el ámbito personal?

Frente a un entorno cada vez más automatizado, tú también tienes **tres poderes**: el poder de conocer lo que ocurre, de exigir explicaciones y de reclamar cuando algo no es justo. No necesitas ser especialista en tecnología para empezar. Solo necesitas informarte, aprender a hacer las preguntas correctas y no conformarte con la primera respuesta.

Con esta guía empieza tu viaje para recuperar el control y hacer valer tus derechos en la era de la IA.



3

Poderes

El derecho a saber, entender y actuar

🕒 10 min

p. 8—16

El derecho a saber

En qué consiste

Entendemos el derecho a saber como un *prederecho*: un derecho previo y complementario al derecho a entender que garantiza algo básico: **conocer cuándo un sistema automatizado está mediando o interviniendo en un proceso**. A menudo, está tan integrada que ni pensamos en ella.

Con qué conecta

Este derecho busca sacar a la luz lo que ahora ocurre en la sombra: el derecho a saber en el ámbito de la IA exige que se informe a las personas si están siendo evaluadas, observadas o tratadas por sistemas automatizados, y bajo qué condiciones. Sin este derecho previo, el derecho a entender pierde sentido: no se puede comprender lo que ni siquiera se sabe.

Cómo pasar a la acción (y qué exigir):

01



Saber que la IA está en funcionamiento en un proceso: pedir que nos informen de forma transparente cuando estamos interactuando con un sistema algorítmico, ya sea de forma directa, como en una conversación con un chatbot, o indirecta, como en una decisión automatizada que te afecta.

02



- a. Cuál es el propósito del sistema automatizado (filtrar CVs, detectar fraude, predecir abandono escolar, segmentar usuarios...)
- b. Qué decisión o acción concreta se toma con esa información
- c. Quién lo está utilizando: si es una empresa, una administración pública, un tercero externo o un proveedor tecnológico.

03



Saber qué datos tuyos se usan, dónde se han obtenido, con qué finalidad se recogieron, durante cuánto tiempo se conservan y si se han compartido o pueden compartir con terceros. Esto incluye también tu derecho a oponerte, corregir errores o solicitar que se dejen de usar en algunos casos: es tu derecho de oposición y limitación del tratamiento de datos.

Cómo pasar a la acción (y qué exigir):

04 ↓

Saber qué persona o entidad tiene la responsabilidad final del uso del sistema y cómo contactar con ella.

05 ↓

Conocer el posible impacto social, como la perpetuación de sesgos y marginación de colectivos vulnerables, las nuevas formas de discriminación, el refuerzo de estereotipos, la contaminación, el consumo energético e hídrico...

El impacto medioambiental de la IA suele quedarse fuera del radar, pero es cada vez más importante, sobre todo en relación a modelos de alto consumo energético e hídrico, como los que hacen posible la IA generativa. La información sobre este impacto es difícil de encontrar, y se basa sobre todo en estimaciones, en torno al **uso tanto de agua y recursos hídricos, como de electricidad** (por ejemplo, para el aire acondicionado que refrigera servidores en grandes centros de datos).

El derecho a entender

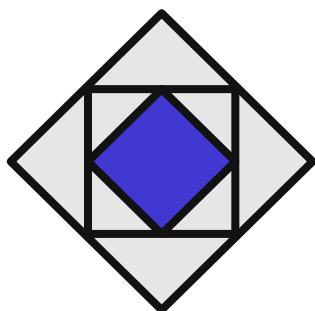
En qué consiste

En un mundo saturado de ruido, el derecho a entender se ha convertido en un derecho fundamental.

Es el derecho de las personas a recibir una Comunicación Clara, a entender la comunicación unidireccional de las Administraciones Públicas y las empresas; a comprender textos legales, jurídicos y financieros; a saber qué significan los términos y condiciones, las políticas de privacidad y otros documentos que impactan nuestra vida cotidiana. Aplicar la Comunicación Clara es la única manera de garantizar este derecho, y va más allá de la creación de documentos fáciles de entender. Consiste en situar al usuario en el centro en todos los pasos y aplicar esta metodología desde el nacimiento de un trámite, producto o servicio, asegurando el lenguaje claro pero también la accesibilidad, la usabilidad, el diseño y la experiencia de usuario.



El cambio hacia una Comunicación Clara implica un compromiso por parte de las organizaciones, desde la administración pública hasta las empresas privadas. Por eso, desde Prodigioso Volcán impulsamos un nuevo Objetivo de Desarrollo Sostenible, el [ODS18](#), que lucha por una comunicación clara, ética y responsable, que asegure el derecho a entender de la ciudadanía con especial atención en el mundo digital.



Ha llegado el día en que también tenemos derecho a esa comunicación clara en el ámbito de la IA, a entender cómo funcionan los sistemas automatizados y a ir más allá. Porque las instituciones y empresas tienen obligaciones de información algorítmica.

Por eso, en el contexto de la IA definimos el derecho a entender como el derecho de toda persona a comprender de forma clara cómo funciona un sistema de IA, en qué momento y con qué propósito está interactuando con él (directa o indirectamente) y cómo se ha llegado a un resultado que le afecta. Es el derecho a acceder a explicaciones comprensibles y contextualizadas sobre el funcionamiento de los sistemas automatizados, sus decisiones y sus limitaciones.

Con qué conecta

Este derecho entronca directamente con dos conceptos, la explicabilidad y la **interpretabilidad** de la inteligencia artificial, que deben abordarse desde la ingeniería y desde el propio proceso de construcción de las soluciones.

- La **explicabilidad** es la cualidad de los modelos de IA para proporcionar una explicación clara y trazable de las razones por las que han llegado a un determinado resultado y qué factores han contribuido al mismo, especialmente en caso de ser adverso para la persona afectada.
- La **interpretabilidad** es la capacidad de la IA de presentar sus procesos y resultados en términos comprensibles para las personas. Incluye también la verificación de que el sistema se comporta de forma esperada para evitar errores, discriminación y sesgos no intencionados.

Cómo pasar a la acción (y qué exigir):

01

Preguntar para entender de manera teórica el funcionamiento del sistema, qué datos utiliza, cómo los procesa y qué lógica sigue para tomar decisiones o hacer recomendaciones, mediante explicaciones claras y adaptadas a diferentes nivel de comprensión.

02

Incidir en los resultados individuales cuando una recomendación derivada de un sistema automatizado afecta sustancialmente a una persona (por ejemplo, en procesos de selección, concesión de ayudas o de créditos, diagnósticos médicos, etc.). Conocer cómo se ha llegado a ese resultado, qué factores han influido en él y si existen vías de revisión, corrección, reclamación, etc.

03

Indagar en las limitaciones del sistema, sus riesgos y aquellos contextos en los que sus predicciones pueden no ser fiables.

04

Examinar la garantía de supervisión humana, especialmente cuando el impacto puede tener consecuencias relevantes.

05

Solicitar una trazabilidad en la construcción del modelo y el flujo de los datos para que especialistas independientes puedan auditarlos.

La obligación de transparencia algorítmica, formulada por primera vez en la Ley Rider española (de 2021), fue una medida pionera en Europa. El artículo introduce la obligación para las empresas de informar a los trabajadores sobre los sistemas de IA que les afectan, en concreto cuando pueden incidir en el acceso, mantenimiento y condiciones de trabajo. En particular, deben especificar los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de IA implicados.

El derecho a actuar

En qué consiste

Una vez que sabemos y entendemos, ya estamos preparados para actuar. Igual que vemos el derecho a saber como un *prederecho* en relación con el derecho a entender, el derecho a actuar sería un *postderecho*: una vez que una persona ha sido informada adecuadamente y ha comprendido cómo el sistema de IA ha llegado a un resultado determinado, debe estar en condiciones de responder, impugnar o influir en ese proceso.

Con qué conecta

El derecho a actuar es la facultad para intervenir de manera efectiva cuando un sistema de IA te afecta de forma negativa, injusta o poco transparente, y asegura algo muy importante: **que las personas no sean tratadas como meros sujetos pasivos.**

Cómo pasar a la acción (y qué exigir):

01

Solicitar la rectificación algorítmica ante cualquier persona o grupo afectado por un tratamiento discriminatorio en el marco de procesos automatizados. Implica la corrección (para que no se vuelva a repetir) y la posibilidad de reclamar compensaciones.

02

Solicitar una revisión humana de decisiones tomadas por sistemas automatizados.

03

Exigir auditorías externas del sistema, y acceder a información sobre los mecanismos de supervisión aplicados.

04

Promover cambios regulatorios o institucionales, con canales efectivos para proponer reformas que permitan un desarrollo más justo, inclusivo y democrático de la IA.

Frente a las injusticias algorítmicas, una persona particular puede encontrar fuerza y respaldo en la colectividad. La participación en asociaciones de consumidores y de defensa de derechos digitales, y en organizaciones de la sociedad civil, así como la colaboración con investigaciones periodísticas, permite visibilizar casos de sesgo, opacidad o uso indebido de sistemas de IA.



¿A quién acudir?

Estas organizaciones (como las 17 que componen el colectivo [IA Ciudadana](#)) ofrecen recursos legales, técnicos y comunicativos para articular demandas colectivas ante instituciones u organismos de transparencia.

El poder de la defensa algorítmica: batallas ganadas

🕒 10 min

p. 17—25

BATALLAS

El reconocimiento y ejercicio de los derechos a saber, entender y actuar frente a los sistemas de IA han dado fruto en muchas ocasiones. La defensa algorítmica ha demostrado tener un impacto real en la protección de derechos fundamentales y en la transformación de sistemas automatizados que eran opacos, erróneos o discriminatorios.

**Hay numerosos
ejemplos de batallas
ganadas, algunas de
ellas en España**

GANADAS

📍 España

Veripol

La “máquina de la verdad” policial



El sistema

En 2018, el Ministerio del Interior presentó Veripol, un sistema de IA para detectar denuncias falsas de robos y hurtos mediante el análisis con IA del lenguaje utilizado en las declaraciones.

El conflicto

Un informe de la Universidad de Valencia reveló deficiencias metodológicas graves, como la asunción directa de que más de la mitad de las denuncias eran falsas, una muestra de entrenamiento insuficiente, de apenas 1000 denuncias, o el hecho de que analizaba los textos redactados por los propios agentes y no las declaraciones textuales de los denunciantes.

La resolución

Veripol estuvo operativo hasta 2024. Tras varios años de implantación opaca en comisarías de todo el país, de solicitudes de transparencia por parte de la academia y de la sociedad civil y de presión mediática, **el Ministerio del Interior anunció su retirada alegando que carecía de validez legal** en procedimientos judiciales.

📍 España

VioGén

La (des)protección
de las víctimas de
violencia de género



El sistema

VioGén es un algoritmo que el Ministerio del Interior puso en marcha en 2007 para evaluar el riesgo de las víctimas de violencia de género de volver a sufrir malos tratos y para activar, en función de ello, las medidas de protección necesarias.

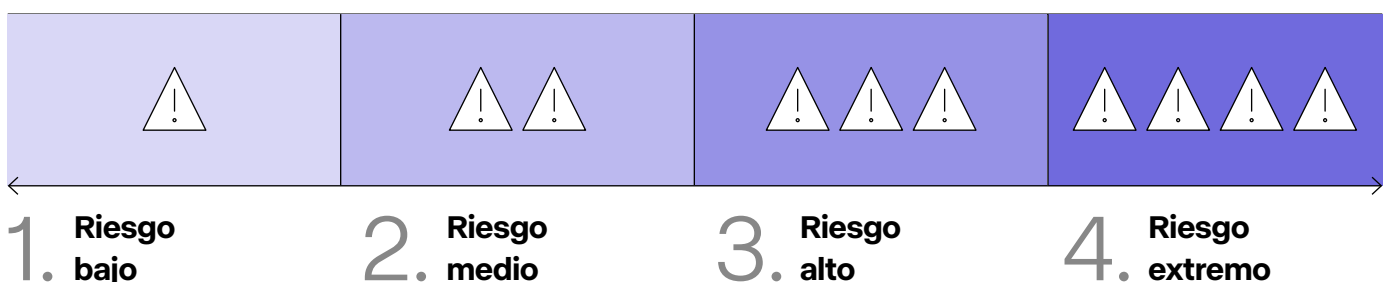
El conflicto

Tras años ofreciendo al Ministerio una auditoría gratuita y confidencial del sistema, la auditora Eticas buscó una forma de hacerla sin su colaboración, gracias al trabajo conjunto con la Fundación Ana Bella. La auditoría reveló problemas graves: principalmente, un 45% de las evaluaciones de riesgo etiquetadas como “no valoradas”, con consecuencias evidentes en las 71 mujeres asesinadas entre 2007 y 2021 que habían denunciado previamente y que no recibieron la protección adecuada. Se detectó también la tendencia del sistema a asignar puntuaciones de bajo riesgo; la imposibilidad de adaptar los cuestionarios a contextos individuales, y la ausencia de corrección de las decisiones algorítmicas por parte de la policía.



La resolución

Tras la publicación de la auditoría de Eticas y un artículo del *New York Times* al respecto, en 2025 se inició el rediseño del modelo. En su nueva versión, entre otros cambios, se establecen cuatro niveles de riesgo y se elimina la opción “no valorada”, para que **ninguna mujer que ha denunciado se clasifique como exenta de riesgo.**



📍 España

El caso de la Ley Rider

el jefe algorítmico



El sistema

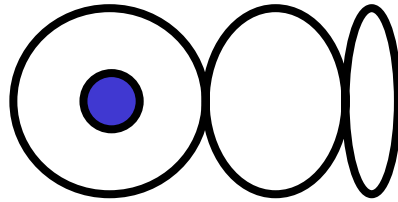
Las plataformas de reparto de comida a domicilio como Glovo, Deliveroo o Uber Eats utilizan algoritmos para asignar pedidos, evaluar el rendimiento y gestionar turnos y tarifas de los repartidores. Estas decisiones automatizadas y tomadas directamente por la IA influyen directamente en sus ingresos, horarios y acceso a futuros encargos.

El conflicto

Este uso opaco y abusivo de los algoritmos ha sido objeto de múltiples denuncias y litigios en España, porque las decisiones del sistema afectaban de forma directa a condiciones laborales sin ofrecer explicaciones ni vías de reclamación.

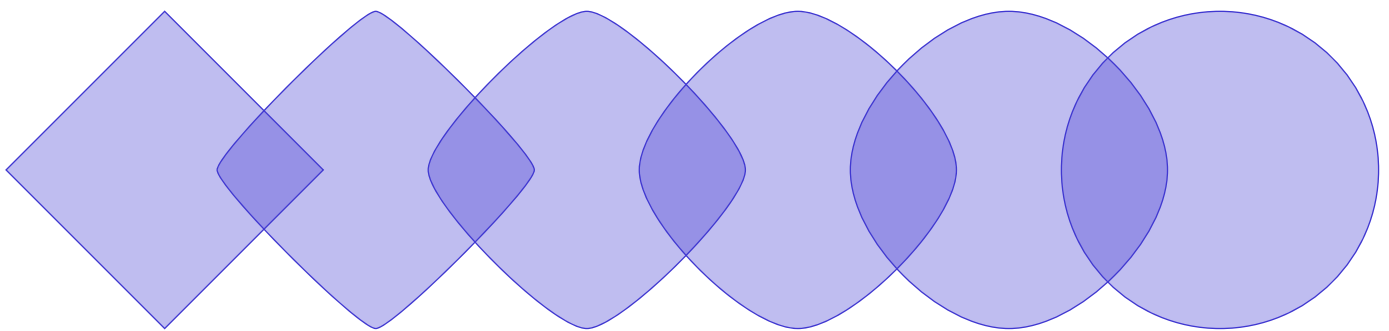
La resolución

En 2021, la Ley Rider reconoció explícitamente el derecho de los trabajadores a conocer cómo funcionan los algoritmos que afectan su trabajo, salario y condiciones laborales.



Esta ley, pionera en Europa, obliga a las empresas a informar a los comités de empresa sobre los parámetros en los que se basan los algoritmos y sistemas de IA que pueden afectar a los trabajadores y a las condiciones de trabajo, y a ofrecer explicaciones por determinadas decisiones automatizadas.

En 2024, la Comisión Europea aprobó una norma parecida: la directiva relativa a la mejora de las condiciones laborales en el trabajo en plataformas.



← **Estructura opaca y abusiva** **Estructura transparente** →

📍 España

Vigilancia universitaria para no copiar

El sistema

Durante la pandemia, varias universidades españolas implementaron sistemas inteligentes de vigilancia para supervisar exámenes online. Estos sistemas usaban tecnologías de IA para el reconocimiento facial, la detección de movimiento y el análisis del comportamiento, con el fin de detectar trampas en los exámenes.

El conflicto

Los estudiantes denunciaron discriminación, falsos positivos y falta de transparencia sobre su funcionamiento. La Agencia Española de Protección de Datos (AEPD) intervino en varios casos y recordó que estas tecnologías deben cumplir con principios de proporcionalidad, necesidad y minimización del tratamiento.

La resolución

Varias universidades retiraron o modificaron su uso. Además, este 2025, la AEPD ha sancionado a la Universitat Internacional de Valencia por el uso obligatorio de este tipo de sistemas, al considerar que el tratamiento de datos biométricos no cuenta con una base legal válida según el Reglamento General de Protección de Datos (RGPD). También desde 2025 está prohibido desplegar y usar sistemas de reconocimiento de emociones basados en IA tanto en los lugares de trabajo como en las instituciones educativas, según la Ley de IA europea.



En Europa, la lista de casos es larga y creciente. Tal vez uno de los más paradigmáticos fue el caso del sistema de puntuación de fraude social SyRI en Países Bajos. El algoritmo se usaba para identificar posibles fraudes en el acceso a ayudas sociales, sin explicar sus criterios ni permitir su cuestionamiento.

📍 Unión Europea

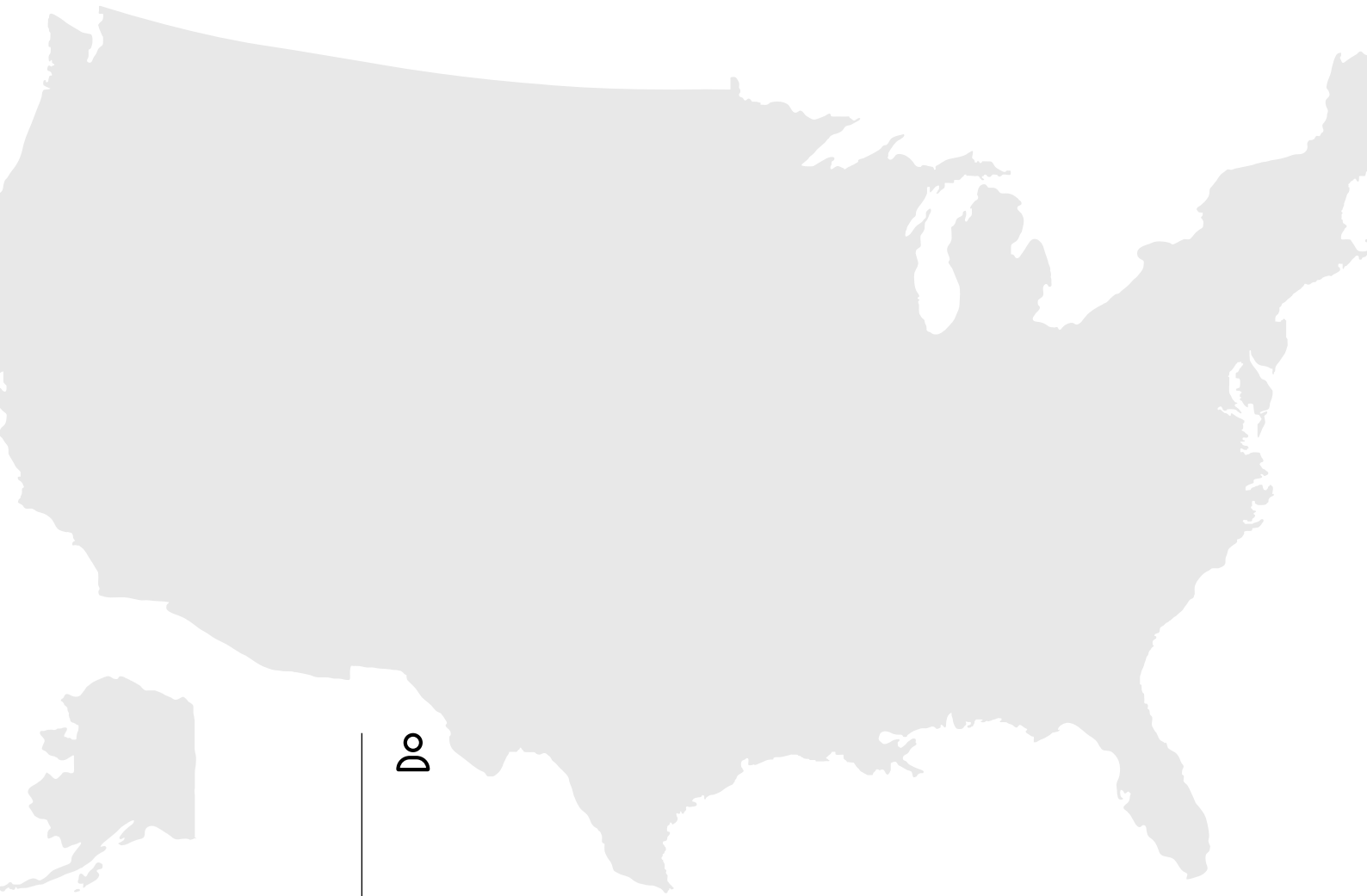
¿Qué pasó con SyRI?

En 2020 fue declarado ilegal por vulnerar derechos fundamentales como la privacidad y la igualdad: en concreto, el artículo 8 del Convenio Europeo de Derechos Humanos (CEDH), que protege el derecho al respeto de la vida privada y familiar. El fallo subrayó la **falta de transparencia, la discriminación indirecta y la imposibilidad de defensa de las personas afectadas**.

Este caso es un hito europeo en defensa algorítmica, y ha tenido consecuencias legislativas a nivel nacional y europeo sobre el uso de IA en servicios públicos.

📍 Estados Unidos

También en Estados Unidos se acumulan casos similares. Lo que comparten todos ellos es el papel fundamental que han tenido las denuncias de personas individuales, así como el trabajo de académicos, auditoras, organizaciones de la sociedad civil y periodistas de investigación, quienes han sacado a la luz fallos graves en sistemas de IA y han impulsado su reforma o retirada.



Do

El ejercicio individual o colectivo del derecho a saber, entender y actuar puede tener un impacto tangible en la protección de las personas y a la vez en la mejora del desarrollo, despliegue y uso de la IA.

¿Qué exige la ley?

🕒 5 min

p. 26—28

¿Qué exige la ley?

Normativas europeas como el RGPD o la Ley de IA establecen obligaciones en materia de transparencia e incluyen aspectos relacionados con los derechos a saber, entender y actuar.

Derecho	Lo que ya sabíamos	Lo nuevo
Derecho a saber	1. El RGPD contempla el derecho a saber qué datos personales se están utilizando, con qué finalidad, de dónde se han obtenido, durante cuánto tiempo se conservarán y si se compartirán con terceros. 2. También garantiza el derecho de acceso, rectificación de errores, oposición al tratamiento, portabilidad y supresión de los datos en determinadas circunstancias. 3. Y además, obliga a que las organizaciones informen claramente sobre quién es el responsable del tratamiento de los datos y cómo contactar con él. 4. En el ámbito del trabajo, la Ley Rider (Ley 12/2021, de 28 de septiembre) obliga a las empresas a informar a los trabajadores sobre los sistemas de IA que les afectan, en particular cuando pueden incidir en el acceso, mantenimiento y condiciones de trabajo.	• La Ley de IA establece que los sistemas de IA que interactúen con los usuarios deben ser identificados como tal. • El contenido generado o modificado con ayuda de IA (imágenes, archivos de audio o vídeo) debe estar claramente etiquetado como “generado por IA”. • Los proveedores de sistemas de alto riesgo deben proporcionar documentación técnica para demostrar requisitos de supervisión humana, precisión y robustez. • Los creadores de modelos de IA generativa deben publicar resúmenes de los datos protegidos por derechos de autor que han utilizado para el entrenamiento. • Cualquier uso protegido por derechos de autor requiere la autorización de su titular. • Los proveedores de sistemas de IA deben registrarlos en una base de datos pública gestionada por la Comisión Europea.([*])

(*) Varias organizaciones de la sociedad civil han criticado las “preocupantes deficiencias” de la base de datos de sistemas de IA que ha creado la Comisión Europea. Entre ellas, que no está claro si entran sistemas no de IA, como algoritmos más simples, y que solo se incluyan obligaciones para los sistemas de alto riesgo, lo que dejaría fuera los chatbots, por ejemplo.

Derecho	Lo que ya sabíamos	Lo nuevo
Derecho a entender	El RGPD establece que cualquier información dirigida al interesado debe ser “concisa, transparente, inteligible y de fácil acceso, con un lenguaje claro y sencillo”.	La Ley de IA establece que las personas afectadas por una decisión de un sistema de IA de alto riesgo con impacto negativo en su salud, seguridad o derechos fundamentales, tendrán derecho a obtener explicaciones claras sobre la función del sistema de IA en el proceso de toma de decisiones y los principales elementos de la decisión adoptada.
Derecho a actuar	1. El RGPD establece que “todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar”. 2. La Ley Rider, por su parte establece que las empresas deben especificar los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de IA que pueden afectar a los trabajadores y a las condiciones de trabajo, y a ofrecer explicaciones por determinadas decisiones automatizadas. 3. También contempla el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión. 4. Existen canales específicos de reclamación, entre los que destacan los habilitados por la AEPD para cuestiones relacionadas con el Reglamento General de Protección de Datos.	1. Se pueden denunciar violaciones de la Ley de IA ante la Agencia Española de Supervisión de la Inteligencia Artificial (AESIA). 2. La Ley de IA habla del derecho a la supervisión humana, que obliga a desarrollar e implementar los sistemas automatizados de modo que una persona humana pueda: • Comprenderlo y supervisarlo para detectar y abordar anomalías; • Interpretar correctamente los resultados; • Decidir no utilizarlo; • Ignorar, anular o revertir su resultado; • Ser conscientes de la tendencia a confiar en exceso en dichos resultados (sesgo de automatización); • E intervenir en su funcionamiento o detenerlo de forma segura.

La coalición IA Ciudadana presentó en 2025 una propuesta para diseñar e implementar un Registro Central de Algoritmos a nivel nacional que sea realmente efectivo y útil, y que incluya tanto los sistemas automatizados utilizados en la Administración Pública como en el sector privado, más allá de los requisitos de la UE.

Las empresas: aliadas clave

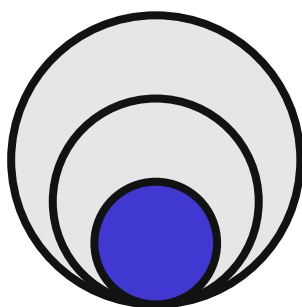
🕒 5 min

p. 29—31

Las empresas: aliadas clave

En paralelo, hay otros movimientos y actores tratando de impulsar y reforzar aspectos relacionados con los derechos a saber, a entender y actuar en el ámbito de la IA. Un aliado clave para ello es el sector privado.

Las empresas que desarrollan, implementan o utilizan sistemas de IA tienen un papel fundamental en la construcción de un ecosistema tecnológico más justo, transparente y respetuoso con los derechos de las personas. Pueden -y deben- ir más allá del cumplimiento de la regulación, y adoptar compromisos relacionados con la ética algorítmica, la explicabilidad y la rendición de cuentas. Porque no es solo una cuestión de responsabilidad social o reputación: también es una cuestión de competitividad y de legitimidad. Las organizaciones que apuesten por una IA más transparente y responsable tendrán mejores productos y estarán mejor posicionadas en un mercado donde la confianza de los usuarios, clientes e inversores es imprescindible.



Las empresas son fundamentales para que la IA del futuro se construya y despliegue desde la responsabilidad, la innovación y la voluntad de hacer las cosas bien.

¿Cómo pueden las empresas ser parte de la solución?



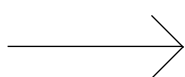
9 acciones

1. **Diseñando sistemas que sean éticos** desde el diseño.
2. **Con transparencia desde el inicio:** incorporando explicabilidad, trazabilidad y supervisión humana en todas las fases de desarrollo e implementación.
3. **Informando de forma clara y accesible** sobre qué sistemas se usan, cómo funcionan y cómo los consumidores, clientes y/o personas afectadas pueden ejercer sus derechos.
4. **Facilitando vías efectivas de reclamación y revisión humana** más allá de simples formularios, con procesos accesibles y acompañamiento real.
5. **Evaluando de forma continua los riesgos** de los sistemas, los sesgos que pueden provocar y su comportamiento a partir de las interacciones de las personas.
6. **Permitiendo auditorías externas independientes** y publicando informes de impacto, tanto técnico como ético, sobre los sistemas desplegados.
7. **Evaluando la evolución de los modelos algorítmicos en el tiempo.**
8. **Participando en la creación de un registro central y público de algoritmos**, contribuyendo a la transparencia colectiva.
9. **Colaborando con la sociedad civil, la academia y la administraciones públicas** en iniciativas conjuntas de evaluación, codiseño, investigación y mejora continua de los sistemas.



7 herramientas prácticas

1. **Evaluaciones de impacto algorítmico:** informes que evalúan los posibles efectos sociales, éticos y legales de un sistema de IA antes de su implementación, incluyendo análisis de riesgos para grupos vulnerables o impactos desproporcionados.
2. **Model cards y system cards:** documentos que resumen información clave sobre un modelo o sistema de IA, como su propósito, datos de entrenamiento, limitaciones, riesgos conocidos y resultados de pruebas.
3. **Red teaming:** ejercicios donde equipos internos o externos simulan usos indebidos del sistema, escenarios imprevistos o ataques, con el fin de identificar vulnerabilidades técnicas, éticas o sociales antes de su despliegue.
4. **Auditorías:** revisiones sistemáticas de los sistemas de IA, tanto antes como durante su despliegue. Pueden enfocarse desde el punto de vista ético, técnico o legal.
5. **Alineamiento:** procesos para garantizar que los modelos se entrenen con objetivos compatibles con los valores humanos para minimizar comportamientos indebidos.
6. **Plataformas automatizadas para monitorizar riesgos:** haciendo un seguimiento en tiempo real del funcionamiento de los sistemas para identificar desviaciones, sesgos emergentes o efectos adversos.
7. **Revisiones éticas:** mediante comités multidisciplinares de especialistas.



Integrar estas prácticas no solo mejora la calidad y seguridad de los sistemas, sino que también facilita la confianza de los consumidores y clientes, y reduce riesgos regulatorios, reputacionales y sociales.

Hacia una defensa algorítmica más activa

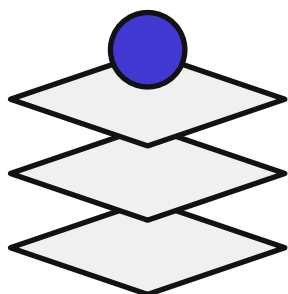
🕒 5 min

p. 32—33

Hacia una defensa algorítmica más activa

Este Manual de defensa algorítmica es una reivindicación y guía para el ejercicio de nuestros derechos frente a la IA y los sistemas automatizados de toma de decisiones. En concreto, nos hemos centrado en tres derechos clave: el derecho a saber, el derecho a entender y el derecho a actuar.

Hemos visto cómo cada uno de estos derechos se traduce en herramientas concretas para no quedarnos al margen de decisiones automatizadas que pueden cambiar el rumbo de nuestra vida: desde saber cuándo y cómo se usa un sistema de IA, hasta comprender su lógica y reclamar ante resultados injustos.



Los casos reales analizados son batallas ganadas que muestran que ejercer estos derechos no solo es posible: es imprescindible. Es lo que permite desmontar sistemas injustos, proteger vidas y definir los límites de lo que consideramos aceptable.

Pero este camino no se puede recorrer en soledad. **La defensa de nuestros derechos en la era algorítmica es una tarea colectiva que requiere vigilancia activa, colaboración y una ciudadanía informada.** Implica exigir a las instituciones que cumplan su función protectora, pero también comprometer a las empresas en la creación de tecnologías más transparentes y auditables.

No se trata de “poner puertas al campo”, sino de construir un futuro digital próspero para todos, donde la equidad, la justicia y la rendición de cuentas no sean opcionales. La posibilidad de preguntar, entender y cambiar las cosas es lo que sostiene una sociedad verdaderamente democrática.

Referencias

- Agencia Española de Protección de Datos. (2024). Conoce tus derechos: Derecho de oposición. <https://www.aepd.es/derechos-y-deberes/conoce-tus-derechos/derecho-de-oposicion>
- Agencia Española de Protección de Datos. (2024). *Ejercicio de derechos*. <https://www.aepd.es/derechos-y-deberes/ejerce-tus-derechos>
- Agencia Española de Protección de Datos. (2020). *Informe sobre reconocimiento facial en exámenes*. <https://www.aepd.es/prensa-y-comunicacion/notas-de-prensa/aepd-publica-informe-reconocimiento-facial-examenes>
- Agencia Española de Protección de Datos. (03 de junio de 2025). *Resolución sobre tratamiento de datos biométricos con IA*. <https://www.aepd.es/informes-y-resoluciones/criterios-juridicos-aepd/aepd-sanciona-tratamiento-datos-biometricos-ia>
- Agencia Española de Protección de Datos. (s.f.). *Sede electrónica: Trámites para ciudadanos*. <https://sedeaepd.gob.es/sede-electronica-web/vistas/infoSede/tramitesCiudadano.jsf>
- García, T., Torrecillas, C., Maqueda, A., Cabo, D. (19 marzo de 2025). *La Policía Nacional deja de usar VeriPol, su IA estrella para detectar denuncias falsas*. Civio. <https://civio.es/transparencia/2025/03/19/la-policia-nacional-deja-de-usar-veripol-su-ia-estrella-para-detectar-denuncias-falsas/>
- Fundación Éticas. (2022). *Auditoría externa del sistema VioGén*. <https://eticasfoundation.org/wp-content/uploads/2025/02/ETICAS-Auditoria-Externa-del-sistema-VioGen--20220308.docx-2.pdf>
- Fundación Éticas. (2025). *The adversarial audit of VioGén, three years later*. <https://eticas.ai/the-adversarial-audit-of-viogen-three-years-later/>
- Gobierno de España. (15 de enero de 2025). *Marlaska y Redondo presentan las mejoras del sistema VioGén*. <https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/paginas/2025/150125-marlaska-redondo-sistema-viogen-2.aspx>

Referencias

- IA Ciudadana. [2025]. Cómo lograr una transparencia real con los registros de algoritmos. http://iaciudadana.org/wp-content/uploads/2025/03/Informe_ES.pdf
- IA Ciudadana. [2025]. *Registros de algoritmos para una transparencia real* (Policy brief). <https://iaciudadana.org/wp-content/uploads/2025/03/Policy-Brief-ESP.pdf>
- Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales. Boletín Oficial del Estado, 233, de 29/09/2021. <https://www.boe.es/eli/es/l/2021/09/28/12/con>
- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Boletín Oficial del Estado, 294, de 06/12/2018. <https://www.boe.es/eli/es/lo/2018/12/05/3/con>
- Ministerio del Interior. [s.f.]. *Sistema VioGén*. <https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen/>
- Satariano, A., Toll Pifarré, R. [17 de enero de 2025]. *Spain Overhauls Domestic Violence System After Criticism*. The New York Times. <https://www.nytimes.com/2025/01/17/technology/spain-domestic-violence-algorithm.html>
- Tascón, M., Montolío, E. [2020]. El derecho a entender. Catarata. <https://comunicacionclara.com/libro-derecho-a-entender/>
- Unión Europea. [2024]. *Reglamento de inteligencia artificial (AI Act)*. <https://artificialintelligenceact.eu>
- Unión Europea. [2024]. *Texto completo del Reglamento de IA*. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202402831 y <https://www.boe.es/doue/2024/1689/L00001-00144.pdf>

En Prodigioso Volcán estamos convencidos de que no hay transformación digital legítima sin la garantía de una protección efectiva de los derechos digitales. Defendemos la tecnología en su expresión más innovadora y brillante, pero siempre dentro del marco irrenunciable del derecho a entender de las personas.



Prodigioso Volcán