

Los sesgos algorítmicos a la luz de los Reglamentos de Inteligencia Artificial y de Protección de Datos

Señala el Parlamento Europeo en un reciente documento sobre la discriminación algorítmica en el marco de los Reglamentos de Protección de Datos (RGPD) e Inteligencia Artificial (RIA)¹ que la diferente consideración regulatoria de los datos personales en ambas normas puede generar inseguridad jurídica y requerir una reforma legislativa o una mayor orientación.

El objetivo del Reglamento de la IA es promover su desarrollo de modo fiable y sostenible, respetando al mismo tiempo los derechos y libertades fundamentales de las personas entre los que se incluye la protección de sus datos personales. Se busca, en este sentido, es mitigar la discriminación y los sesgos en el desarrollo, la implementación y el uso de sistemas de IA de alto riesgo, para lo cual se permite el tratamiento de categorías especiales de datos personales, con base en un conjunto de condiciones (por ejemplo, medidas de protección de la privacidad) diseñadas para identificar y evitar la discriminación que pudiera producirse al utilizar estas nuevas tecnologías. Sin embargo, el RGPD parece más restrictivo en este aspecto.

En septiembre de 2024, la autoridad de supervisión belga señaló en un informe² cómo la corrección del sesgo en los datos de entrenamiento para los sistemas de IA es coherente tanto con el principio de «tratamiento justo» de datos establecido en el RGPD como con la necesidad de garantizar la exactitud de la información personal recopilada por los encargados del tratamiento de datos.

Algunos sistemas de IA pueden causar discriminación y vulnerar los derechos fundamentales de las personas. Por ejemplo, los sistemas de IA generativa no se clasifican de alto riesgo como tales, pero en algunos casos pueden representar amenazas para los derechos y libertades fundamentales, como cuando generan discursos de odio. Algunos sistemas de IA generativa que vulneran los derechos fundamentales podrían incluirse en la lista de sistemas de alto riesgo que los reguladores pueden actualizar periódicamente, según lo establecido en el Reglamento. La discriminación también podría surgir en el caso de los vehículos autónomos; por ejemplo, si se desarrollan de forma que detecten con mayor precisión a los peatones de piel más clara que a los de piel más oscura.

Los algoritmos de selección de candidatos a un puesto de trabajo pueden contener sesgos, por ejemplo, en relación con el género o la salud.

¹[https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/769509/EPRS_ATA\(2025\)769509_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/769509/EPRS_ATA(2025)769509_EN.pdf)

² <https://www.autoriteprotectiondonnees.be/publications/artificial-intelligence-systems-and-the-gdpr---a-data-protection-perspective.pdf>

Los sistemas de IA se utilizan cada vez más en los sectores bancario y crediticio como herramienta de apoyo para evaluar la concesión de préstamos o hipotecas. Estos sistemas pueden seguir procesos de toma de decisiones que ocultan la discriminación o enmascaran sesgos basados en factores como la residencia o la etnia de los clientes.

Con el fin de garantizar la detección y corrección de sesgos en los sistemas de IA, el artículo 10(5) del Reglamento permite el tratamiento de categorías especiales de datos personales «en la medida en que... sea estrictamente necesario a efectos de garantizar la monitorización, detección y corrección de sesgos en relación con los sistemas de IA de alto riesgo», y siempre que se garanticen las garantías adecuadas para el respeto de los derechos fundamentales. Este enfoque clasifica los sistemas de IA según el grado de riesgo que representan para los derechos y libertades fundamentales de las personas. Al determinar si un sistema se califica como de alto riesgo (artículo 6 del RIA), se hace hincapié en el alcance del sistema y la tarea específica que desempeña.

El RGPD, por su parte, define en su artículo 9 el concepto de categorías especiales de datos personales: origen racial o étnico, opiniones políticas, creencias religiosas o filosóficas, afiliación sindical, datos genéticos, datos biométricos, datos de salud, prácticas u orientación sexual. Algunos académicos han señalado que, para detectar si un algoritmo utilizado en el sector laboral discrimina por motivos étnicos, en principio, la organización en cuestión necesita conocer la etnia de los solicitantes de empleo.

Artículo 10(5) del RIA se enmarca dentro del ámbito de aplicación del artículo 9 del RGPD, aclarando que, en caso de conflicto, prevalece el RGPD, que no debe verse afectado en su aplicación. Para tratar categorías especiales de datos personales, deben respetarse los fundamentos específicos enumerados en el artículo 9 del RGPD (por ejemplo, el consentimiento explícito del interesado).

Algunos expertos consideran que el tratamiento de datos personales en virtud del artículo 10 del RIA también podrían permitirse en virtud del "motivo sustancial de interés público" previsto en el RGPD; de hecho, su artículo 9.2.g) contempla el tratamiento de categorías especiales de datos cuando sea necesario por razones de interés público sustancial, con base en el Derecho de la Unión o de los Estados miembros, que será proporcional al objetivo perseguido, respetará la esencia del derecho a la protección de datos y establecerá medidas adecuadas y específicas para salvaguardar los derechos fundamentales y los intereses del interesado.

En todo caso, Para que el artículo 10.5 del RIA se interprete plenamente conforme con el RGPD, deben cumplirse algunas condiciones:

El tratamiento de categorías especiales de datos personales debe llevarse a cabo con sólidas medidas de ciberseguridad para evitar fugas de datos, tal como exige el RGPD;

Los sistemas de IA que tratan datos personales deben cumplir con principios del RGPD como la minimización de datos, la limitación de la finalidad y el almacenamiento, la integridad y la confidencialidad, y la privacidad desde el diseño y por defecto;

El requisito de necesidad, principio clave del RGPD, debe concebirse como la autorización para que las entidades procesen categorías especiales de datos solo cuando sea estrictamente necesario para proteger otros derechos fundamentales, como la no discriminación;

Para que el tratamiento de datos sea lícito, deben concurrir los fundamentos establecidos en el artículo 9 del RGPD para el tratamiento de datos sensibles.

Además de las categorías especiales de datos enumeradas en el artículo 9 del RGPD, puede producirse discriminación como resultado del tratamiento de datos personales que revelen otros factores, como la edad o el sexo. Estas no son categorías especiales de datos según la legislación de la UE, sino simplemente datos personales. Por lo tanto, la base jurídica para el tratamiento de estos datos será la que se enumera en el artículo 6 del RGPD.

El artículo 6 del RGPD tiene un contenido más amplio que el artículo 9 del RGPD, ya que contempla un mayor número de bases jurídicas para el tratamiento de datos personales (por ejemplo, el interés legítimo). Esto significa que, cuando surge discriminación con respecto a factores que no se consideran categorías especiales de datos personales, el tratamiento de datos para mitigar el sesgo puede realizarse de conformidad con los fundamentos del Artículo 6, que son más laxos y fáciles de probar que el interés público o el consentimiento expreso.

En términos generales, parece prevalecer una incertidumbre compartida sobre cómo debe interpretarse la disposición de la Ley de IA sobre el tratamiento de categorías especiales de datos personales para evitar la discriminación. El RGPD, que impone límites al tratamiento de categorías especiales de datos personales, podría resultar restrictivo en un contexto dominado por el uso de la IA en muchos sectores de la economía y ante el tratamiento masivo de datos personales y no personales. Una reforma del RGPD o nuevas directrices sobre su interacción con la Ley de IA podrían ayudar a abordar estas cuestiones.