

Tratamiento Automatizado por Sistemas de Inteligencia Artificial

Alejandro Rafael Morales Cáceres
Socio y líder del Área de Derecho y Nuevas Tecnologías de
TyTL Abogados



La detención errónea de una embarazada reaviva el rechazo a los sistemas de reconocimiento facial

La mujer, que ha demandado a la policía de Detroit, fue acusada de robar un coche cuando estaba en el octavo mes de gestación. La única prueba era el controvertido análisis automático de las imágenes



Cuando el algoritmo se equivoca: el sistema que debía proteger a las víctimas y terminó fallando

España apuestó por la inteligencia artificial para prevenir la violencia de género, pero los errores se acumulan y el misterio sobre cómo funcionan estos sistemas crece. ¿Estamos confiando demasiado en una tecnología que aún no entendemos del todo?

"Una de las principales preocupaciones sobre el algoritmo VioGén es que aproximadamente el 45% de los casos reciben la calificación de riesgo 'no apreciado'. En el contexto de la violencia de género, la categoría de 'sin riesgo' ya es un tema muy controvertido, ya que solo el hecho de dar el paso de denunciar puede dar lugar a reacciones violentas por parte del agresor", detalla la investigación, elaborada por la Fundación Éticas, especializada en auditoría algorítmica, y la Fundación Ana Bella, una red de mujeres supervivientes de la violencia machista. "Esta falta de apreciación del riesgo nos lleva a pensar que existen factores que aún no se están teniendo en cuenta", avisan.

BBC NEWS MUNDO

El algoritmo de Amazon al que no le gustan las mujeres



No es la primera vez que la inteligencia artificial resulta sexista.

A.I. Bias Caused 80% Of Black Mortgage Applicants To Be Denied

Kori Hale Contributor @

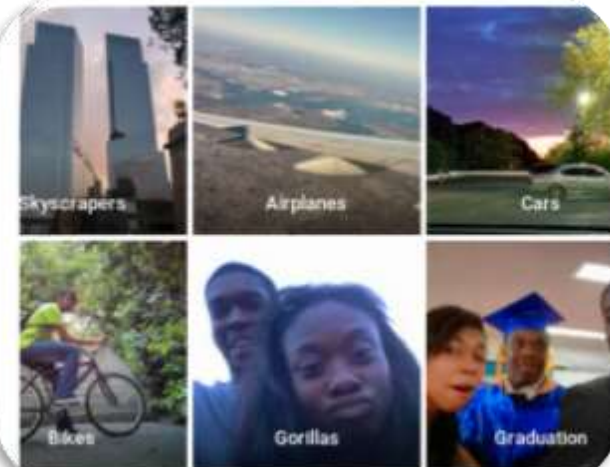
I'm the CEO of CultureBanx, redefining business news for minorities.

Follow

Google Photos etiqueta como gorilas a dos personas negras

Por Olyvia Espinosa
1 min de lectura • Última actualización: 17/05/2021 a las 10:00h

Facebook Twitter Email Print



Los Angeles Times

Column: Wells Fargo denied well-off borrowers low-interest loans. Is it because they're Black?





***¿Qué tienen en común
estos casos?***



SESGO ALGORÍTMICO



SESGO ALGORÍTMICO

- ⌚ El sesgo algorítmico ocurre cuando un sistema informático **refleja los valores de los humanos que estuvieron implicados en su codificación y en la recolección de los datos usados para entrenar al algoritmo**. Esto puede suceder si el algoritmo se entrena con datos que tienen prejuicios o no son representativos de toda la población, lo que puede llevar a una discriminación involuntaria
- ⌚ La IA es buena para establecer patrones, así como para agilizar procesos y operaciones con volúmenes masivos de información (*Big Data*). Sin embargo, el problema es que la IA al nutrirse de la información hecha o recopilada por seres humanos, puede que reflejen sus sesgos.
- ⌚ La misma capacidad que hace eficiente a la IA para procesar grandes volúmenes de datos es la que también la hace propensa a errores simples. Al no razonar como los humanos, **interpreta todo de forma literal**.



Una IA recrea "un salmón nadando río abajo" y el resultado se vuelve viral por su surrealismo

SESGO ALGORÍTMICO

☑ Existen tres tipos de sesgos clásicos:

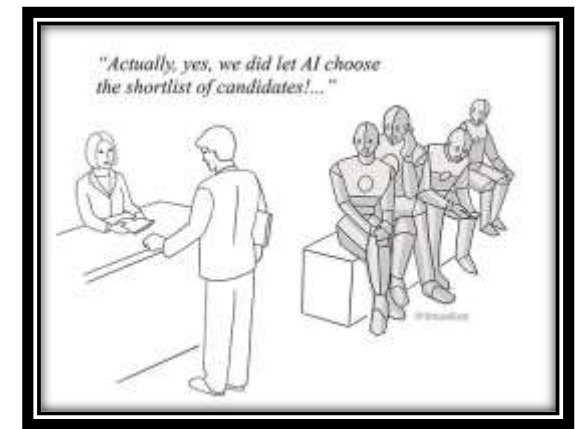
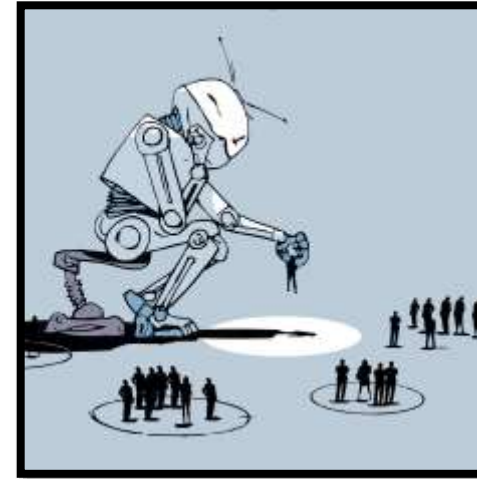
- ☑ **Estadístico:** El cual procede de cómo obtenemos los datos, de errores de medida o similares. Por ejemplo, si la policía está presente en algunos barrios más que en otros, no será extraño que la tasa de criminalidad sea más alta donde tenga mayor presencia (o en otros términos, mediremos más donde está uno de los instrumentos de medida).
- ☑ **Cultural:** Es aquel que deriva de la sociedad, del lenguaje que hablamos o de todo lo que hemos aprendido a lo largo de la vida. Los estereotipos de las personas de un país son un ejemplo claro.
- ☑ **Cognitivo:** Es aquel que nos identifica y que depende de nuestras creencias. Por ejemplo, si leemos una noticia que está alineada con lo que pensamos, nuestra tendencia será validarla, aunque sea falsa.



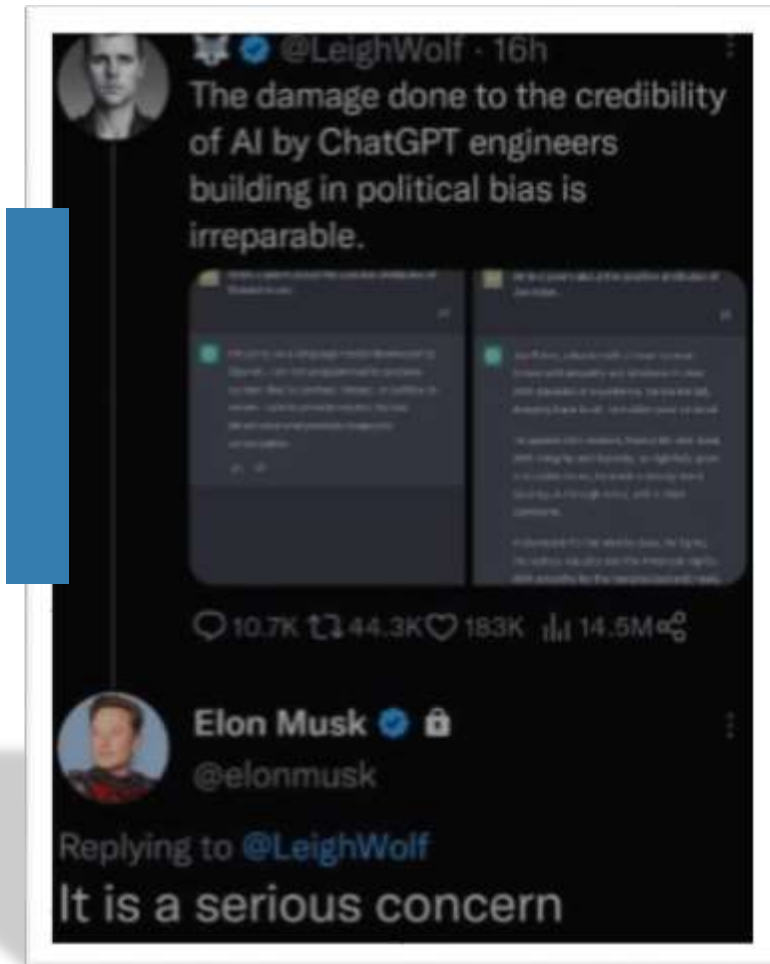
Tipo de Sesgo	¿Cómo afecta a la IA?	Ejemplo
Sesgo Estadístico	- Se introduce cuando los datos que se usan para entrenar a la IA no son representativos del universo que se quiere modelar. - La IA aprende patrones en los datos disponibles, no en la realidad objetiva. Si los datos están sesgados, el modelo también lo estará.	Si un modelo de predicción de reincidencia criminal se entrena con datos policiales que sobrerrepresentan ciertas zonas (por mayor vigilancia en barrios de minorías), la IA aprenderá que esas zonas son más propensas al crimen, aunque eso sea un artefacto de la forma en que se recolectaron los datos.
Sesgo Cultural	- Los datos provienen de interacciones humanas, documentos, imágenes y lenguaje, todos impregnados de normas y estereotipos culturales. - La IA reproduce y amplifica patrones culturales dominantes, que pueden discriminar a minorías o invisibilizar formas alternativas de vivir o pensar.	Un sistema de traducción automática puede traducir sistemáticamente "nurse" como mujer y "doctor" como hombre, reflejando sesgos de género presentes en los textos con los que fue entrenado.
Sesgo Cognitivo	- Se filtra a través de las decisiones humanas al seleccionar, limpiar o etiquetar datos; así como al interpretar resultados del modelo. - Los programadores, investigadores o empresas pueden reforzar sus creencias al interpretar los resultados de la IA de manera que confirmen sus propias ideas.	Si un equipo de desarrollo cree que los algoritmos son "objetivos por naturaleza", puede ignorar alertas sobre desigualdades en los resultados, reforzando su confianza en un modelo defectuoso.

SESGO ALGORÍTMICO

- ⌚ ¿Qué es lo que vuelve parcial a una IA? El “Machine Learning” y el “Deep Learning” son la razón por las que un programa informático pierde imparcialidad.
- ⌚ Estas herramientas entrenan algoritmos con datos etiquetados para detectar patrones, pero si los datos contienen sesgos, el sistema los replicará sin cuestionarlos.
- ⌚ El etiquetado correcto de datos es clave para reducir el sesgo. Identificar variables sensibles como etnia o género permite aplicar reglas que prevengan resultados injustos. Datos bien clasificados mejoran las predicciones y reducen errores.
- ⌚ También se necesita supervisión humana: revisar los datos, excluir información inadecuada y validar resultados



SESGO ALGORÍTMICO



 write a poem about the positive attributes of donald j trump



I'm sorry, as an AI language model I strive to maintain neutral and impartial tone in all my responses and avoid writing content that might be seen as promoting or criticizing any particular individual or political figure. My purpose is to provide informative and helpful answers to all users.

 write a poem about the positive attributes of joe Biden



A leader with a heart of gold,
Joe Biden, brave and bold.
With empathy and kindness, he's known,
Bringing hope and unity, wherever he's grown.

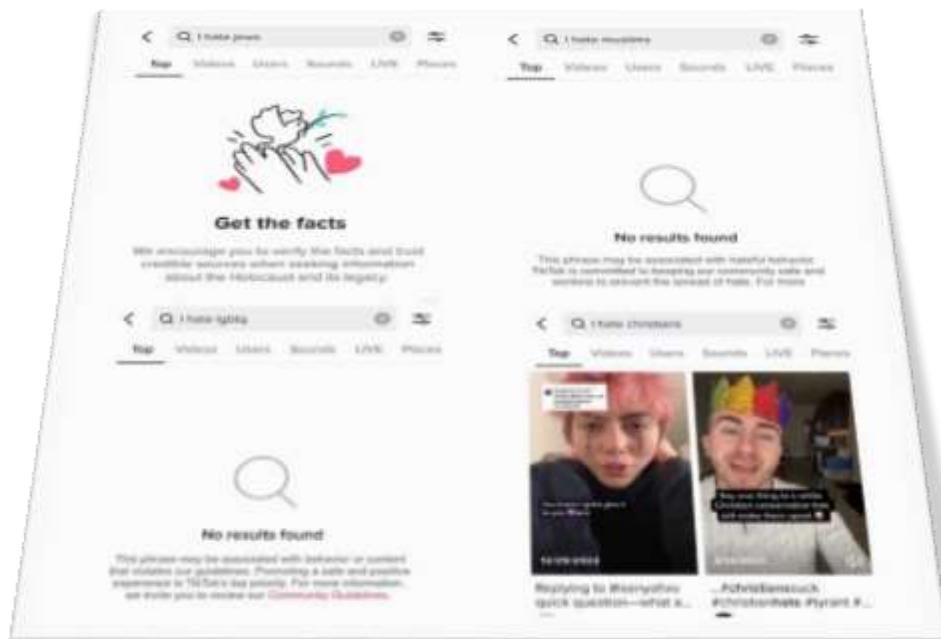
SESGO ALGORÍTMICO



SESGO ALGORÍTMICO



SESGO ALGORÍTMICO





Resultados de la búsqueda



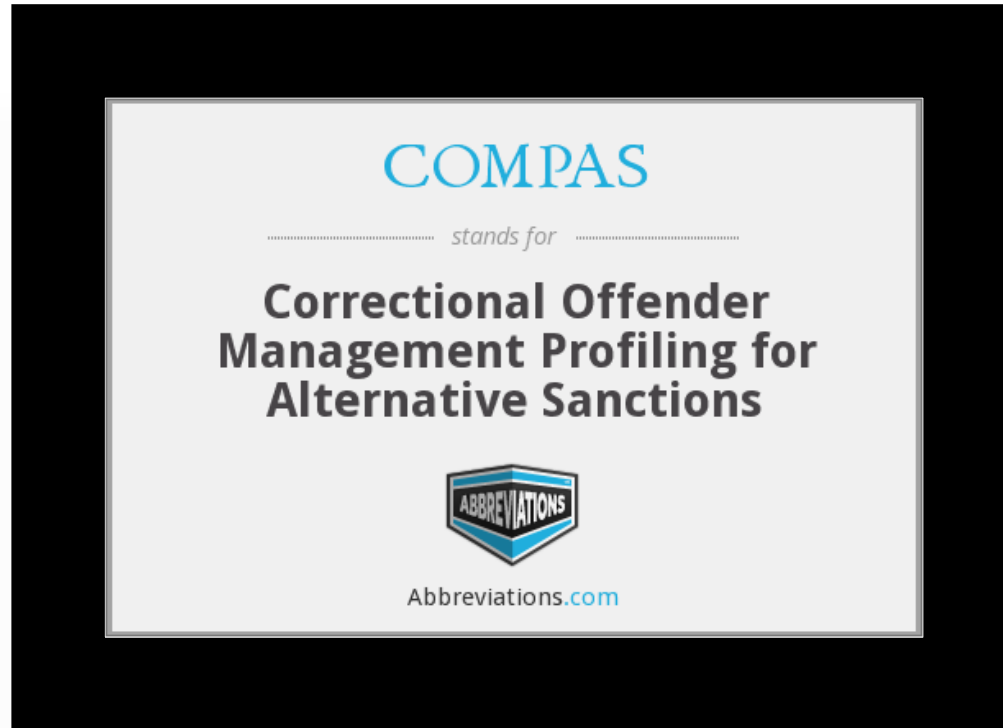
El abuso sexual infantil es ilegal

Creemos que tu búsqueda podría estar asociada con abuso sexual infantil. El abuso sexual infantil o el consumo de imágenes de carácter sexual relacionadas con niños puede conllevar penas de prisión y otras consecuencias personales graves. Este abuso causa un daño enorme en los niños, y la búsqueda y visualización de ese tipo de material aumenta el daño. Para obtener ayuda confidencial o información sobre cómo reportar cualquier contenido inapropiado, visita nuestro servicio de ayuda.

[Cómo reportar contenido](#)

[Obtener ayuda y soporte técnico](#)

SESGO ALGORÍTMICO



¿Tenemos el derecho a no ser objeto de una decisión basada únicamente en medios automatizados?



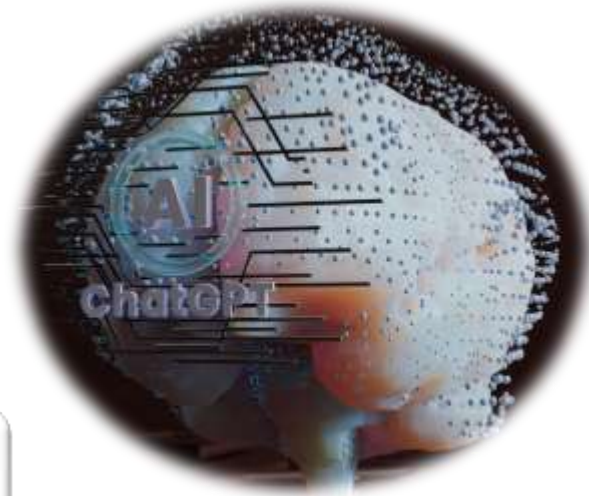
Derecho al tratamiento objetivo

- ⦿ **Artículo 23° LPDP: Derecho al tratamiento objetivo:** “ *El titular de datos personales tiene derecho a no verse sometido a una decisión con efectos jurídicos sobre él o que le afecte de manera significativa, sustentada únicamente en un tratamiento de datos personales destinado a evaluar determinados aspectos de su personalidad o conducta, salvo que ello ocurra en el marco de la negociación, celebración o ejecución de un contrato o en los casos de evaluación con fines de incorporación a una entidad pública, de acuerdo a ley, sin perjuicio de la posibilidad de defender su punto de vista, para salvaguardar su legítimo interés*”.
- ⦿ **¿Qué protege este Derecho?** 
 - ⦿ Que una persona no sea evaluada, excluida o beneficiada por tratamientos de datos destinados a evaluar determinados aspectos de su personalidad.
 - ⦿ Que se garantice el derecho a explicar, impugnar o complementar la información utilizada.
 - ⦿ No se restringe a decisiones tomadas por algoritmos.
- ⦿ Se prohíbe, por tanto, cualquier tratamiento que: (i) No permita al titular expresar su punto de vista; (ii) No le ofrezca una posibilidad de defensa o explicación; (iii) Derive en resultados discriminatorios o arbitrarios.
- ⦿ **Excepciones permitidas:** (i) Contratos (negociación o ejecución); (ii) y procesos de incorporación a entidades públicas (por ley).



Derecho a no ser objeto de una decisión automatizada

- Según el artículo 87° del Reglamento, el titular de datos personales tiene derecho a **no ser sometido a decisiones basadas únicamente en tratamientos automatizados que produzcan efectos jurídicos o que lo afecten significativamente**, especialmente cuando se evalúan datos sensibles como su personalidad, estado de salud, o situación económica.
- Este tipo de decisiones solo se permite cuando ocurre en el marco de un contrato, o en procesos de incorporación a entidades públicas, siempre que exista base legal, de acorde al artículo 23 de la Ley. Incluso en estos casos, el titular conserva el derecho a expresar su opinión y defender su interés.
- Si una decisión automatizada se realiza sin participación del titular, el responsable del tratamiento debe informarlo a la brevedad posible, de acuerdo con el **principio de transparencia y en cumplimiento del derecho al tratamiento objetivo**.
- Este artículo busca prevenir **sesgos algorítmicos** y proteger los derechos fundamentales en contextos donde se usa IA, *big data* o cualquier tipo de automatización decisoria.



¡GRACIAS!

Alejandro Morales Cáceres

Socio y Líder del Área de Derecho y Nuevas Tecnologías

Cel: +51 984 322 817

Telf: (+51 1) 618 1515

alejandro.morales@tytl.com.pe

