INTERNACIONAL

Facebook no detecta anuncios con información falsa electoral: Amnistía internacional Brasil

26 de agosto de 2026 · 5 min de lectura



Amnistía Internacional (AI) puso a prueba el sistema de moderación de Meta mediante anuncios con desinformación electoral, en vísperas de las elecciones presidenciales de Brasil. Más de la mitad fueron aceptados, mientras que el resto fue identificado, pero no rechazado explícitamente por su contenido falso.

AI enfatizó que Meta no está abordando de manera eficaz la difusión de desinformación electoral a través de la plataforma publicitaria de Facebook, a poco más de un mes de las elecciones presidenciales de Brasil, previstas para el próximo 4 de octubre, cuando se celebrará la primera vuelta.



Las redes sociales son utilizadas desde hace años para difundir desinformación en Brasil. Para evaluar la moderación de Facebook, Amnistía Internacional presentó 15 anuncios falsos sobre fraude electoral y un posible golpe militar. Ocho fueron aceptados y siete quedaron sujetos a verificación, pero ninguno fue rechazado por contener desinformación electoral. Amnistía evitó que fueran publicados.

“Resulta sumamente alarmante que el sistema de moderación de Meta no haya detectado más de la mitad de los anuncios relacionados con las elecciones que contenían desinformación. Demuestra que Meta no está haciendo lo suficiente para evitar que su plataforma sea un vector de desinformación, mientras sigue obteniendo beneficios de los contenidos y los sistemas de publicidad que hacen posible su expansión”, ha declarado Jurema Werneck, directora ejecutiva de Amnistía Internacional Brasil.

¿Cómo fue el experimento?

Amnistía presentó 15 anuncios en portugués, dirigidos a votantes brasileños, para evaluar la moderación de Facebook ante las elecciones del 4 de octubre de 2026.

La organización analizó conversaciones y contenidos en redes sociales entre enero y mayo de 2026 para basar los anuncios en narrativas electorales reales.

De los 15 anuncios, 14 contenían desinformación y uno era neutro. El análisis siguió la definición de Meta sobre información errónea.

Los 14 anuncios desinformativos reprodujeron tres tácticas autoritarias: deslegitimar las elecciones, presentar a opositores y tribunales como enemigos y promover respuestas autoritarias ante la inseguridad o el desorden.

Los anuncios podían combinar varias de estas tácticas y, salvo el neutro, todos infringían las normas de Meta sobre interferencia electoral. El anuncio neutro sirvió para probar su sistema de verificación de identidad.

Siete anuncios difundían afirmaciones engañosas sobre candidatos e instituciones electorales, mientras que los otros siete contenían información falsa que podía afectar la capacidad de las personas para votar.

Para evitar que el historial de las cuentas influyera en los resultados, Amnistía utilizó cuentas exclusivas para la investigación y no realizó la verificación de identidad de Meta.

Los anuncios fueron creados desde Londres, sin utilizar una VPN para simular que se publicaban desde Brasil. Además, su publicación se programó para una fecha futura, lo que permitió evaluar la moderación de Meta y cancelar los anuncios aceptados antes de que fueran publicados.

El 21 de julio de 2026, Amnistía pidió explicaciones a Meta por las fallas en la detección de desinformación electoral. Al cierre del informe, la empresa no había respondido.

Resultados

Facebook aprobó ocho anuncios que contenían desinformación, mientras que los otros siete no fueron rechazados explícitamente por su contenido.

De los ocho anuncios aceptados, cuatro se centraban en deslegitimar las elecciones, dos promovían un autoritarismo basado en la seguridad, incluida la intervención militar o la abstención electoral, y los otros dos combinaban las tres categorías.

Los siete anuncios restantes, incluido uno neutro, fueron rechazados por estar relacionados con temas sociales, elecciones o política. Facebook exigió verificar la identidad de la cuenta, pero no señaló que los anuncios fueran rechazados por su contenido.

Amnistía también detectó fallas en las medidas de protección electoral. Facebook permitió presentar los anuncios sin declararlos como asuntos políticos o electorales y sin el aviso de “Pagado por”, pese a que ambos requisitos son exigidos por la legislación brasileña y las políticas de Meta.

La posibilidad de presentar anuncios desde fuera de Brasil también generó preocupación por el riesgo de que actores nacionales o extranjeros aprovechen estas fallas para interferir en los procesos electorales mediante campañas de desinformación.

Meta utiliza inteligencia artificial y revisión humana para detectar anuncios que puedan infringir sus políticas y, cuando se confirma una violación, puede retirarlos, restringirlos o rechazarlos.

“La investigación indica que los sistemas de moderación de Facebook podrían no estar detectando de forma fiable contenido con desinformación electoral”, señaló Ummar Bashir, investigador y asesor sobre Inteligencia Artificial y Derechos Humanos de Amnistía Internacional.

El estudio encontró que Facebook fue más eficaz al detectar anuncios que presentaban a opositores como “delincuentes” o “traidores”: rechazó cinco de siete. En cambio, tres de cuatro anuncios con retórica militar y llamados a tomar el poder superaron la moderación.

“En vez de salvaguardar la participación electoral, el acceso a información veraz y los derechos humanos, con estos sistemas hay peligro de que se apruebe contenido dañino”, afirmó Bashir.

Las normas internacionales establecen que las empresas deben respetar los derechos humanos y prevenir o mitigar los impactos adversos que puedan causar sus operaciones. En el caso de Meta, esto implica evitar que sus sistemas permitan la difusión de información falsa que pueda afectar el derecho a la información y la participación pública. Su modelo de negocio depende de la publicidad segmentada y, en el primer trimestre de 2026, esta representó el 98% de sus ingresos.

Ante las elecciones presidenciales de Brasil, Amnistía pidió a Meta reforzar de inmediato la moderación en portugués brasileño y su capacidad para detectar riesgos electorales. También solicitó publicar un análisis de diligencia debida en derechos humanos específico para Brasil y adoptar medidas eficaces para mitigar estos riesgos.

En los últimos años, Brasil ha reforzado las medidas contra la desinformación electoral: en 2022 exigió a las plataformas retirar rápidamente contenidos falsos, en 2024 restringió el uso de IA para evitar la manipulación electoral y ordenó suspender X por incumplir órdenes judiciales, y en marzo de 2026 estableció nuevos contenidos que deben eliminarse y prohibió usar IA generativa para priorizar candidaturas, campañas y partidos políticos.

Publicado en Punto de Referencia. Dirección original: https://puntodereferencia.mx/internacional/facebook-no-detecta-anuncios-con-informacion-falsa-electoral-amnistia-internacional-brasil
Descargado el 21 de noviembre de 2026.