

Octubre 2025

Boletín mensual de monitorización del discurso de odio en redes sociales

Este boletín tiene como objetivo analizar los contenidos de discurso de odio que circulan por las redes sociales, monitorizados y notificados a las plataformas de internet (Facebook, Instagram, X, TikTok y YouTube) en el mes de octubre de 2025. Así, los datos permiten explorar tanto las características y patrones del discurso como los grupos de población a los que está dirigido. En particular, se analiza la reacción de las plataformas ante la retirada de contenido, los grupos diana objeto del discurso de odio, los episodios prototípicos que lo suscitan y el tipo de contenido y lenguaje utilizado en los mensajes de discurso de odio.

Nota 1: Todos los gráficos y análisis presentados en este boletín fueron elaborados con datos del Sistema FARO (elaboración propia).

Nota 2: Los datos presentados en este boletín deben ser interpretados con cautela, dado que el Sistema Faro se lanzó en marzo de 2025 y todavía está en fase de optimización de la herramienta de inteligencia artificial.

Contenidos monitorizados

Durante el mes de octubre de 2025, el monitor FARO detectó un total de 52.958 contenidos reportables de discurso de odio en redes sociales. Las plataformas retiraron el 42% de los contenidos reportados, lo que supone un incremento de 3 puntos porcentuales (7,7%) en la tasa de retirada respecto al mes anterior, evidenciando una mejora en la efectividad de las acciones de moderación por parte de las plataforma frente a los reportes realizados.

52.958

Mensajes detectados

42%

% Mensajes retirados

Contenidos retirados

Las plataformas han retirado el 9% de los contenidos notificados como usuario normal, lo que supone una disminución de 3 puntos porcentuales respecto al mes anterior (12%). En cuanto al plazo de retirada, el 6% se efectuó dentro de las primeras 24 horas, el 1% en un plazo de 48 horas y el 2% restante durante la primera semana tras la notificación.

Por otra parte, utilizando la vía *trusted flagger*, las plataformas han retirado un 33% de los contenidos reportados, lo que representa un incremento de 5 puntos porcentuales respecto al mes anterior (27%).

Los resultados muestran una diferencia significativa en la eficacia de las vías de notificación de retirada a las plataformas, lo que pone de manifiesto la necesidad de fortalecer los mecanismos de moderación y de mejorar la efectividad de los reportes realizados como usuarios normales, para que estos también se traduzcan en la retirada de los contenidos de discurso de odio.

Actualmente, se trabaja conjuntamente con las plataformas mediante grupos de trabajo y reuniones bilaterales, con el objetivo de optimizar estos procedimientos y reforzar la detección proactiva de contenidos de discurso de odio.

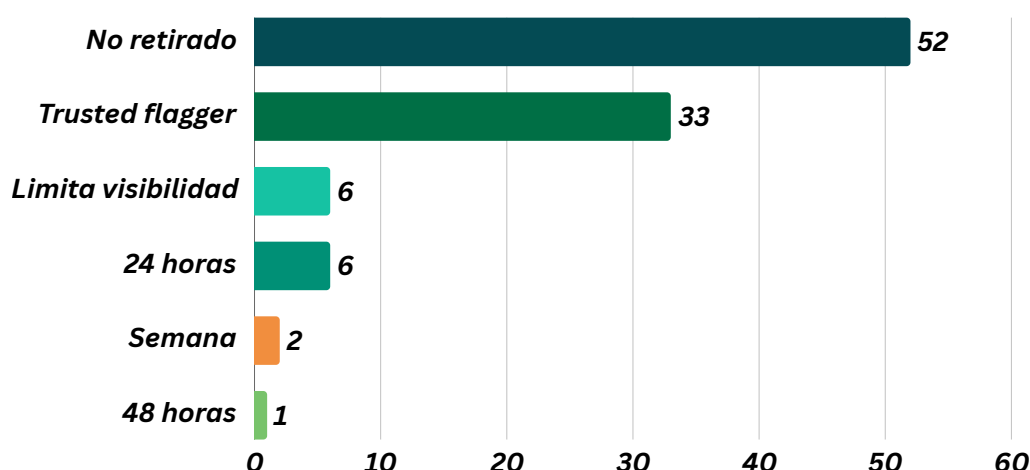
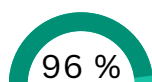
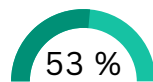


Gráfico reacción de las plataformas ante la retirada de contenido (%)



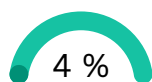
TikTok aumenta en 6 puntos porcentuales su tasa de retirada (96%) respecto al mes anterior (90%), siendo un 27% de los contenidos retirados en las primeras 24 horas.



Facebook ha retirado más de la mitad de los contenidos notificados, en su gran mayoría a través la vía *trusted flagger* (95%).



Instagram retira el 37% de los contenidos. Solo un 5% de los contenidos han sido retirados como usuario normal.



X ha sido la plataforma que menos contenidos ha retirado con solo un 4%. No obstante, el 86% de los contenidos se han retirado en el transcurso de una semana.



YouTube tiene una tasa de retirada del 22%, retirando un 83% de los contenidos utilizando la vía *trusted flagger*.

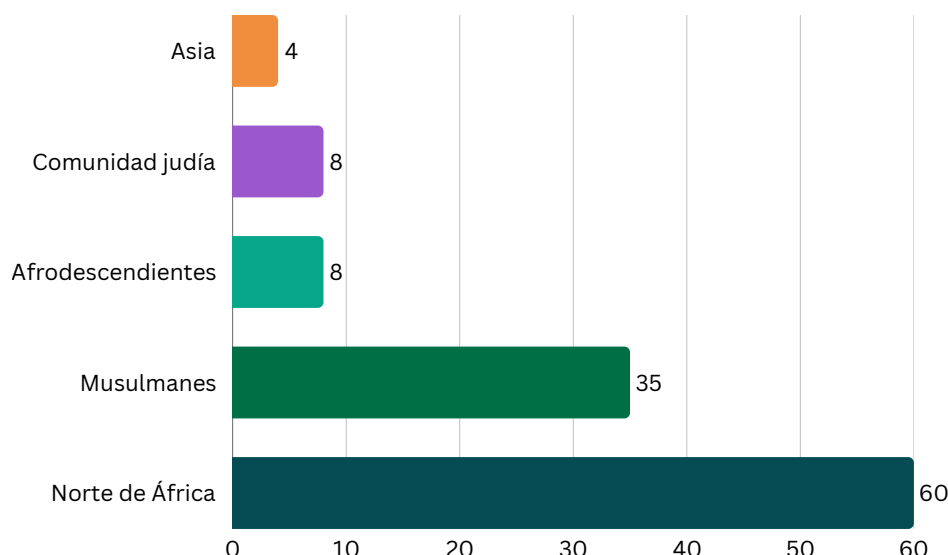
Características del discurso de odio

Grupo diana

En octubre continuó predominando la hostilidad en redes sociales hacia las personas del norte de África, que concentraron el 60% de los contenidos de discurso de odio analizados, sin variaciones respecto al mes anterior.

Es importante destacar que los mensajes de odio dirigidos hacia las personas musulmanas mantienen su tendencia al alza por segundo mes consecutivo, representando el 35 % de los mensajes detectados, 2 puntos porcentuales más que en septiembre (33 %).

Por otra parte, se observó un incremento de 1 punto porcentual en los contenidos de odio hacia la comunidad judía, que alcanzaron el 8 % del total analizado.

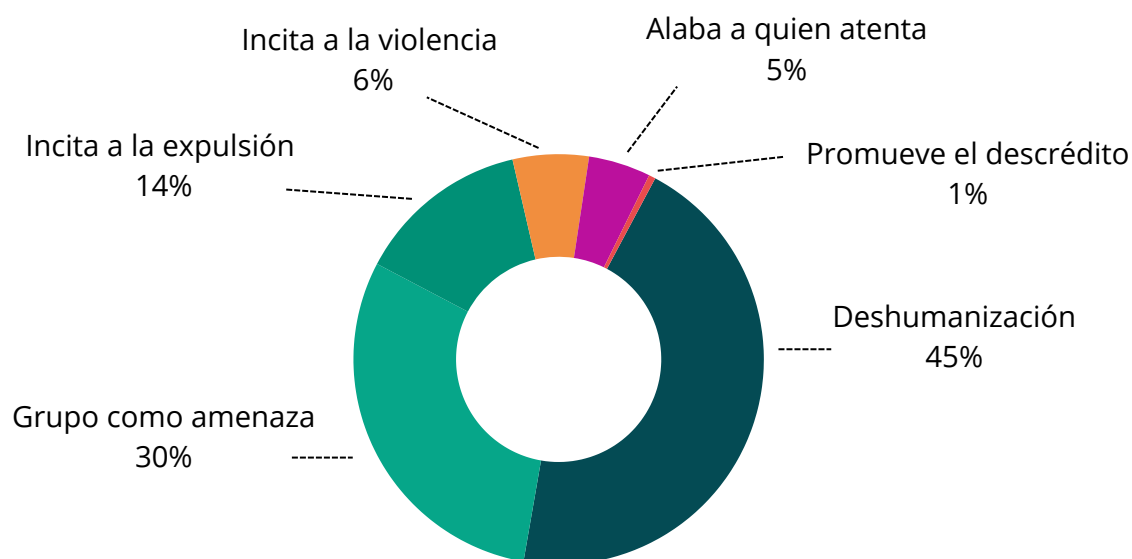


Tipo de contenido

Los contenidos que deshumanizan a personas de origen extranjero aumentaron notablemente respecto a septiembre, alcanzando el 45 % del total, 14 puntos porcentuales más que el mes anterior, y se consolidan como el tipo de contenido más común, lo que refleja el riesgo persistente de normalización de actitudes hostiles hacia estas personas en redes sociales. En contraste, los contenidos que alaban a quien atenta y los que incitan a la violencia disminuyeron considerablemente pasando del 14 % al 5 % y del 13 % al 6%, respectivamente.

Los mensajes que presentan al grupo como una amenaza se mantuvieron relativamente estables (30%), con una reducción de 3 puntos porcentuales respecto a septiembre, aunque siguen contribuyendo a una percepción de peligro sobre las personas de origen extranjero.

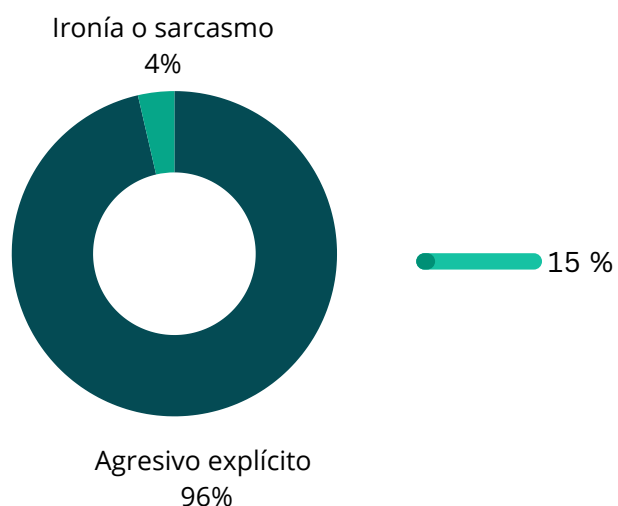
Por su parte, los contenidos que incitan a la expulsión de personas migrantes aumentaron hasta el 14%, lo que representa un incremento de 6 puntos porcentuales en comparación con septiembre, señalando un mayor riesgo de estigmatización y rechazo social.



Expresión del lenguaje

El 96% de los contenidos de discurso de odio utiliza un lenguaje agresivo explícito, caracterizado por palabras o expresiones directas que atacan, humillan o descalifican a las personas de origen extranjero. Además, estos contenidos recurren con frecuencia a estereotipos y etiquetas que reducen a las personas a identidades simplificadas, como en el caso de comentarios del tipo: "moros = ladrones". El lenguaje agresivo explícito se utiliza para expresar hostilidad o rechazo hacia las personas migrantes y su uso se ve favorecido por la sensación de anonimato en internet, con ejemplos de comentarios como: "*a tu puto país moro de mierda*"; "*exterminio*" o "*moros, hijos de puta*".

En cambio, el 4% de los contenidos utiliza la ironía o el sarcasmo para difundir discursos de odio de forma encubierta. Aunque menos frecuente, este tipo de discurso es igualmente nocivo, ya que minimiza la gravedad y favorece su aceptación social. Entre los ejemplos se encuentran: "*Alérgico al jamón*"; "*Moros, que raro...*" o "*pocas pateras se hunden*".



El 15% de los contenidos reportados utiliza emojis, imágenes, videos o lenguaje codificado. Esto refleja una adaptación del odio a los mecanismos de moderación y censura de las plataformas, donde los usuarios recurren a códigos, símbolos o imágenes para transmitir mensajes sin ser detectados por los sistemas automáticos ("*88*"; "*PMDM*" o "*put os mor os*"). Además, el uso de emojis y memes puede servir para enmascarar o disimular la violencia. Con comentarios como: "🔥🐼🔥"; o "*Hay que aniquilar el 🇹🇷*".

Episodios que suscitan discurso de odio

En octubre, los contenidos asociados a la inseguridad ciudadana representaron el 38% del total, lo que supone un aumento de dos puntos porcentuales respecto al mes anterior.

Uno de los principales detonantes de inseguridad ciudadana ha estado relacionado con la noticia sobre la investigación policial del joven detenido por la presunta agresión sexual a una menor en Hortaleza (Madrid), que generó una elevada cantidad de mensajes que incitaban a la violencia o a la expulsión de niños, niñas y adolescentes no acompañados. Entre los comentarios más frecuentes se identifican expresiones como *"castración y deportación"*, *"fuera menas"* o *"fuera invasores parásitos"*.

De manera similar, la viralización de un video en el que una persona del norte de África intenta acceder a una vía por el carril contrario provocó comentarios deshumanizantes y xenófobos, utilizando el término *"moro"* para estigmatizar a las personas migrantes. Con comentarios como: *"moro mierda"* o *"vulgar rata de cloaca"*. Se identificaron también expresiones de tono irónico o sarcástico dirigidas a desacreditar al grupo diana, como: *"enséñenles que invadir no está bien"* o *"cómo no iba a ser moro"*

Asimismo, la noticia de la agresión a un joven con discapacidad por parte de varias personas de origen extranjero, así como la difusión del video en internet, ha generado comentarios relacionados con la inseguridad ciudadana. Los comentarios deshumanizan y degradan a las personas migrantes e incitan a su expulsión aludiendo a una falta de integración. Con comentarios como: *"cerdos inmigrantes"*; *"el 90% de los magrebíes son animales que solo merecen morirse"* o *"deportación inmediata"*.

Estos sucesos muestran cómo la inseguridad ciudadana actúa como un medio para reproducir discursos xenófobos, racistas y de estigmatización de los diferentes grupos diana, generando un impacto que trasciende el ámbito mediático y moldea la percepción social sobre la migración.

Respecto al conflicto armado (41%), las manifestaciones pro palestinas celebradas en Madrid y Barcelona, la cobertura mediática del regreso a España de la *"flotilla"* tras los acontecimientos con Israel y la noticia de la posible retirada de Alemania del evento cultural Eurovisión en caso de excluir a Israel, provocaron que el episodio prototípico *"conflicto armado"* continuara siendo, por segundo mes consecutivo, el principal detonante de discurso de odio durante octubre. No obstante, se observa un descenso de siete puntos porcentuales respecto al mes anterior.

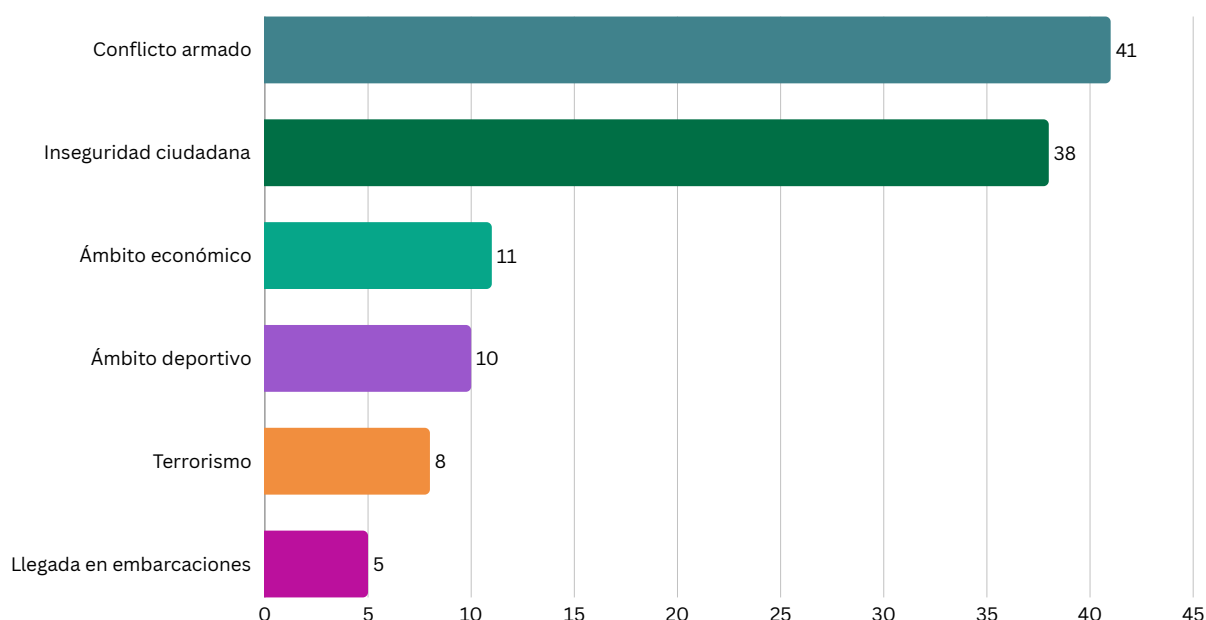
Las noticias relacionadas con el conflicto entre Israel y Palestina generaron mensajes en los que se acusa a la población palestina de ser criminal, junto con comentarios negativos hacia personas judías y musulmanas, tales como: *"lo decía el pintor austriaco, que estos de Israel eran malos..."*, *"los moros son traicioneros"* o *"pobres palestinos terroristas"*.

En el ámbito económico (11%), la noticia del anuncio del Gobierno de que más de 200.000 familias no reúnen los criterios necesarios para acceder al Ingreso Mínimo Vital (IMV) dio lugar a contenidos con discurso de odio. Los comentarios detectados cuestionan el destino de los recursos públicos y señalan a los niños, niñas y adolescentes no acompañados como supuestos beneficiarios de las ayudas estatales, reforzando la narrativa de que las personas migrantes tienen un “beneficio indebido” con mensajes como: “*me hacía pasar por mena para participar de este parasitismo*” o “*el dinero de los impuestos gastándose en mantener moros*”.

Por último, en el ámbito deportivo (10%), el partido de LaLiga disputado el 26 de octubre entre el Real Madrid y el FC Barcelona registró un incremento en la actividad vinculada al discurso de odio. En esta ocasión, los mensajes se concentraron en torno a la figura del jugador Lamine Yamal, objeto de ataques racistas y xenófobos que evidencian cómo el deporte continúa funcionando como un espacio de expresión de tensiones sociales y prejuicios raciales.

Las expresiones detectadas como “*subnormal moro*” o “*el moro mierda*” reflejan una racialización del discurso deportivo, donde la identidad étnica o el origen del jugador se convierten en ejes de hostilidad. Este tipo de comentarios no solo vulnera la dignidad individual, sino que reproduce estereotipos y dinámicas de exclusión que trascienden el ámbito futbolístico, reforzando actitudes discriminatorias en el espacio online y en la esfera pública.

Nota: Por razones de protección de datos y responsabilidad institucional, los ejemplos de mensajes han sido omitidos o adaptados para evitar la reproducción literal de expresiones de odio.



Edita y distribuye: Observatorio Español del Racismo y la Xenofobia (OBERAXE)