

Julio 2026

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 julio de 2026. 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 encuentra en fase de optimización de su herramienta de inteligencia artificial.

Contenidos monitorizados

En el mes de julio de 2026, el monitor FARO detectó un total de 57.406 contenidos de discurso de odio y narrativas discriminatorias en redes sociales.

Sobre la muestra de contenidos de discurso de odio seleccionados y reportados a las plataformas, la tasa de retirada se situó en 70%, lo que supone 9 puntos porcentuales más que el mes anterior (61%).

57.406

Mensajes detectados

70%

% Mensajes retirados

Contenidos retirados

Las plataformas eliminaron el 12% de los contenidos notificados como usuario normal, una proporción ligeramente superior a la observada en junio (10%). En cuanto a los tiempos de respuesta, el 9% de los contenidos se retiró durante las primeras 24 horas, un 1% adicional en las siguientes 48 horas y el 2% restante a lo largo de la primera semana desde su notificación.

En relación con la vía *trusted flagger*, las plataformas retiraron el 58% de los contenidos reportados, lo que representa un incremento de 7 puntos porcentuales respecto al mes anterior (51%).

Los resultados muestran una eficacia considerablemente mayor de la vía *trusted flagger* frente a las notificaciones realizadas por usuarios normales. Para continuar avanzando en la mejora de estos procesos, se mantiene un trabajo conjunto con las plataformas a través del grupo de trabajo sobre discurso de odio en redes sociales, reuniones bilaterales y canales de comunicación directa.

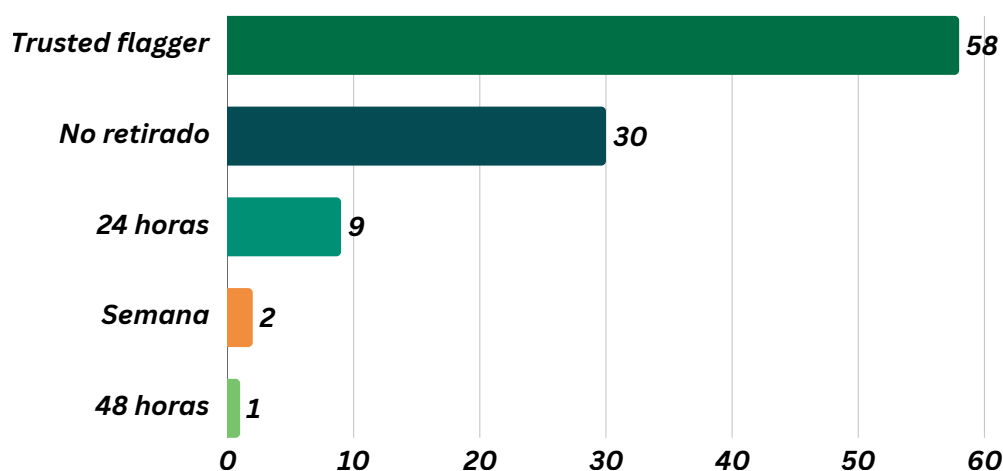
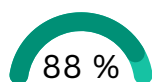
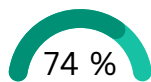


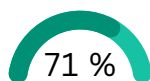
Gráfico reacción de las plataformas ante la retirada de contenido (%)



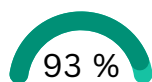
TikTok retiró el 88% de los contenidos notificados, de los cuales el 40% se eliminó en las primeras 24 horas.



Facebook eliminó el 74% de los contenidos notificados, lo que supone una mejora de 12 puntos respecto al mes anterior (62%).



Instagram alcanzó una tasa de retirada del 71%, situándose por encima del doble de la registrada el mes anterior (32%).



X registró una tasa de retirada del 93%, concentrándose el 99% de las retiradas en la vía *trusted flagger*.



YouTube alcanzó una tasa de retirada del 14% sobre los comentarios reportados, correspondiendo el 100% de las retiradas a la vía *trusted flagger*.

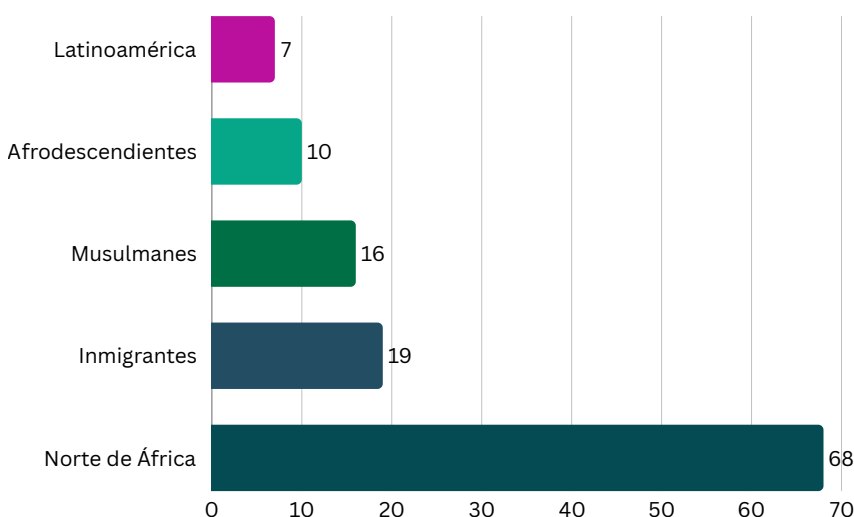
Características del discurso de odio

Grupo diana

Durante el mes de julio, las personas del Norte de África continúan siendo el principal grupo diana del discurso de odio en redes sociales, al concentrar el 68% de los contenidos detectados. Si bien este porcentaje es inferior al registrado en junio (88%), se mantiene en niveles elevados, confirmando la especial incidencia de la hostilidad dirigida hacia este grupo diana.

Por su parte, el contenido dirigido hacia la categoría genérica de personas inmigrantes representó el 19% de los mensajes analizados, una proporción notablemente superior a la observada en meses anteriores. Este aumento coincide con distintos acontecimientos vinculados al ámbito migratorio y las políticas públicas en esta materia, como se detalla en el apartado relativo a los episodios prototípicos.

Por su parte, el grupo diana de personas musulmanas concentra el 16% de los contenidos detectados, 5 puntos porcentuales más que en junio (11%).

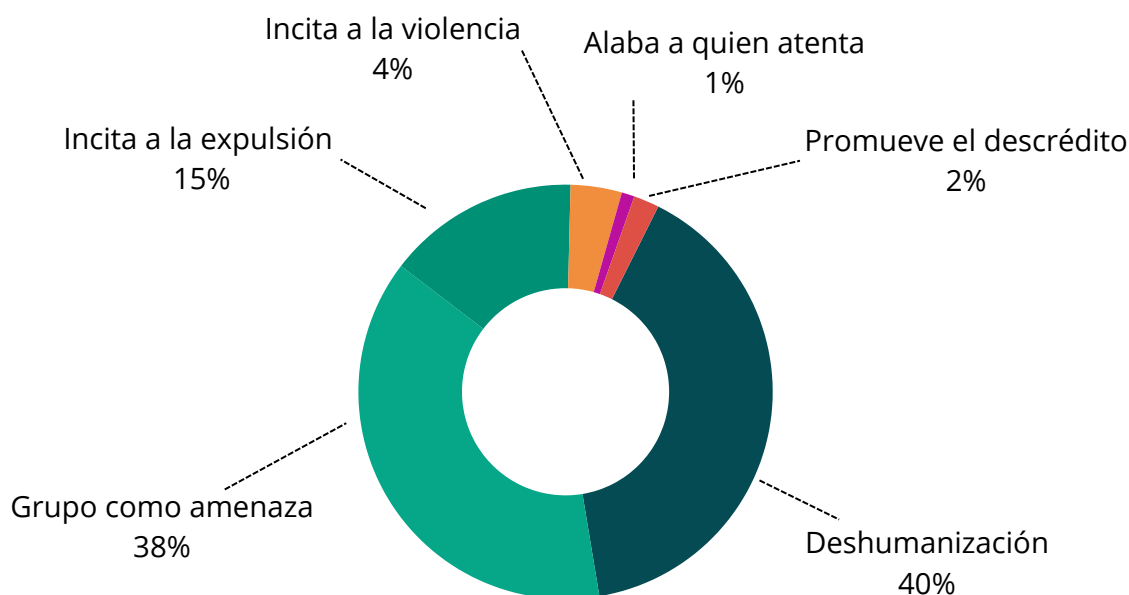


Tipo de contenido

En el mes de julio, los contenidos que deshumanizan o degradan gravemente suponen un 40% del total de contenidos, lo que implica un descenso de 8 puntos porcentuales respecto al mes de junio (48%). En segundo lugar, los contenidos que presentan a los grupos diana como una amenaza representan el 38% del total de contenidos, lo que supone un incremento de 6 puntos porcentuales con respecto al mes anterior (32%).

Por otro lado, el contenido que incita a la expulsión supone un 15% del total de contenido, experimentando un aumento de 5 puntos porcentuales con respecto al mes de junio (10%).

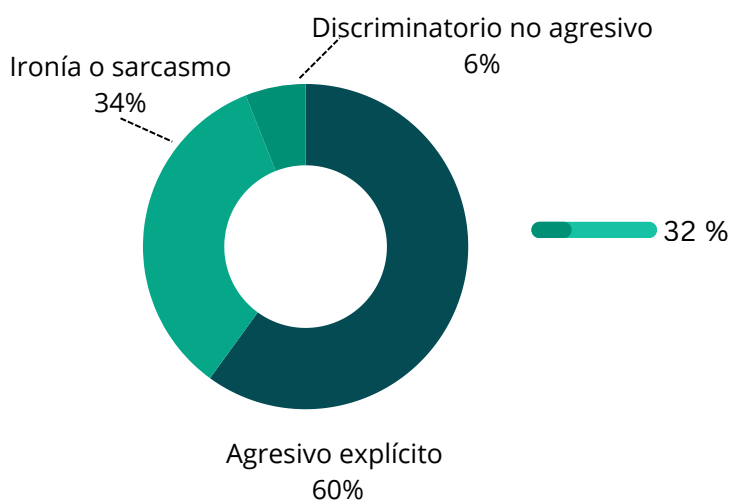
Asimismo, el contenido que incita a la violencia representa un 4% del total, descendiendo 2 puntos porcentuales en relación con el mes anterior. Por último, los mensajes que promueven el descrédito suponen un 2% del total, y los que alaban a quienes atacan contra los grupos diana alcanzan un 1% del total de contenido.



Expresión del lenguaje

El 60% del contenido examinado presenta un lenguaje agresivo explícito, basado en insultos, amenazas y descalificaciones directas dirigidas a los grupos diana, lo que pone de manifiesto un alto grado de hostilidad en el discurso, tal y como se aprecia en expresiones como: “caña al mono”, “la escoria que regularizan 🔥”, “vuelve al estercolero del que vienes”.

Por otra parte, el 34% del contenido analizado recurre al uso de la ironía o el sarcasmo para vehicular narrativas discriminatorias mediante un tono aparentemente humorístico que favorece la normalización de prejuicios y contribuye a la creación de un clima hostil hacia los grupos diana. Este tipo de mensajes supone, además, un mayor desafío para los sistemas de detección automática empleados por las plataformas, con comentarios como “la playa está llena de cucarachas”, “que no aparezcan furgonetas de españoles cansados y a cazar” o “los bajan de los árboles y pasa esto”.



Finalmente, el 33% del contenido analizado incluye imágenes, vídeos, memes o el uso de lenguaje codificado, adoptando un formato de carácter visual. Este tipo de contenidos posee una mayor capacidad de difusión y de impacto y, simultáneamente, dificulta la identificación automática y precoz del discurso de odio por parte de las plataformas.

Algunos ejemplos ilustrativos de este tipo de contenido son: “purga ya ✂️ 🔥 🖊️ 🇪🇸 🇵🇹”, “moro de 💩” o “fuera de nuestros barrios 🗑️ 🗑️”

Episodios que suscitan discurso de odio

Durante el mes de julio, la inseguridad ciudadana constituyó el principal episodio prototípico que generó discurso de odio, concentrando el 48% de los contenidos monitorizados. Los mensajes identificados reprodujeron narrativas de criminalización de las personas migrantes, construidas tanto sobre acontecimientos específicos como sobre contenidos descontextualizados que refuerzan estereotipos negativos y percepciones de amenaza.

En concreto, uno de los acontecimientos que más discurso de odio ha catalizado en redes sociales, fue la llegada de miles de personas procedentes de Marruecos a la Ciudad Autónoma de Ceuta intensificada el día 30 de julio. En este contexto, se han identificado numerosas narrativas deshumanizantes, como por ejemplo *"que asco de raza"*; narrativas que presentan a las personas inmigrantes como una amenaza, con mensajes como *"moros invasores de mierda"*; así como contenido que promueve el descrédito hacia los grupos diana, con comentarios como *"sois lo contrario a civilizados, un español vale por siete moros"*.

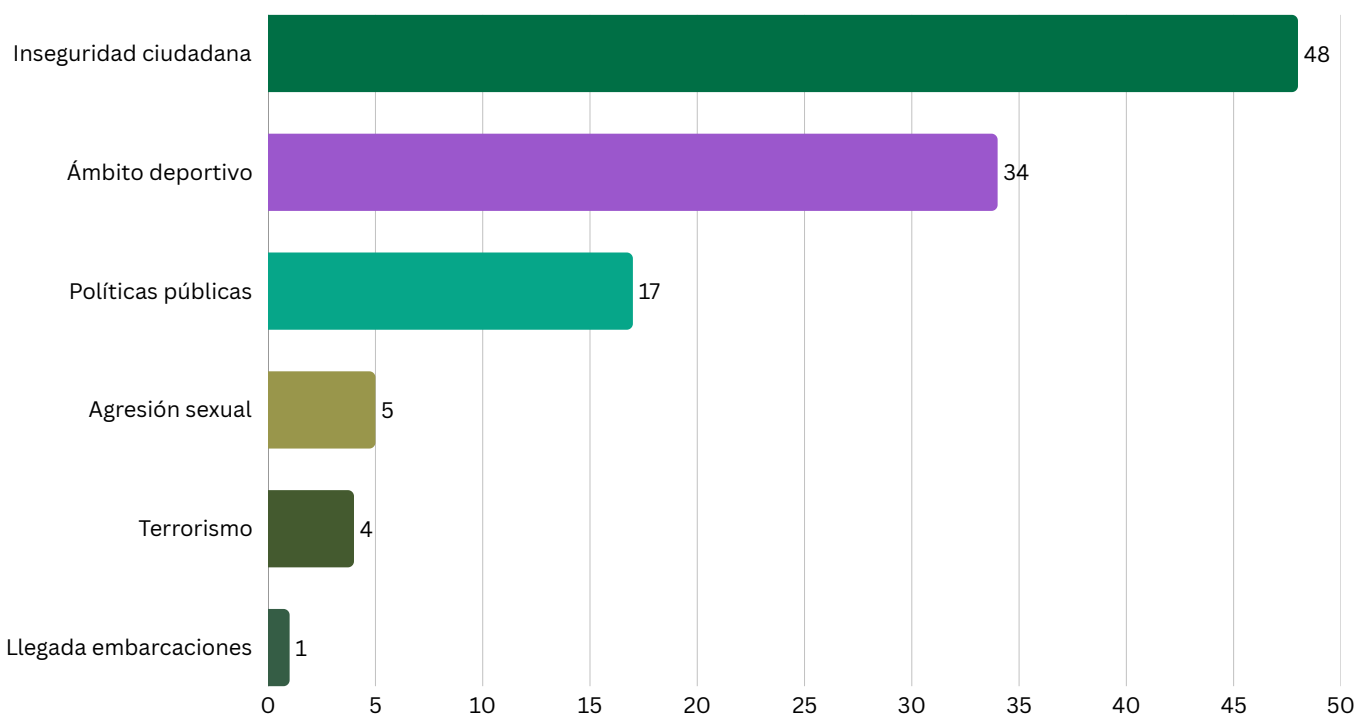
Por su parte, el ámbito deportivo, con un 34% de los mensajes analizados, adquiere una relevancia especialmente destacada en la circulación de discursos de odio. Durante la Copa Mundial de Fútbol de 2026 se detectaron contenidos hostiles contra distintas selecciones nacionales y sus seguidores. Entre los principales focos de hostilidad destacaron los mensajes racistas e islamóforos contra Lamine Yamal y su entorno, con expresiones como *"que deje de jugar el moro basura"*. Por otro lado, los cuestionamientos sobre la composición de determinadas selecciones nacionales, especialmente la francesa, se reflejaron en comentarios como *"no es racismo, no parece Francia"* o *"son todos angoleños, sin el futbol serían sintecho"*. Asimismo, la selección de Marruecos y sus aficionados fueron objeto de comentarios que los presentaban como una amenaza para la convivencia y la seguridad, con mensajes como *"invasión musulmana"* o *"son animales salvajes, que les eliminen ya"*. En torno a la final entre España y Argentina se registró un incremento de los mensajes xenóforos dirigidos contra personas argentinas y latinoamericanas, que se prolongó tras la finalización del Mundial, con expresiones como: *"argensimio de mierda"*, *"os odiamos más después del mundial, sudacas"*.

En cuanto al discurso de odio asociado al episodio prototípico de las políticas públicas, este representó el 17% del total de los contenidos analizados durante el mes de julio. Una parte significativa de los mensajes monitorizados estuvo vinculada a la regularización extraordinaria de personas migrantes, así como a diversas informaciones sobre su legalidad, evolución o sobre los principales países de origen de las personas solicitantes. En este contexto, se detectaron narrativas que presentan los posibles beneficiarios como una amenaza para la convivencia, la seguridad ciudadana y los recursos públicos, reforzando percepciones de rechazo hacia las personas migrantes.

Por otro lado, los contenidos sobre el Pacto de Migración y Asilo de la Unión Europea, así como sobre las demandas de Baleares de mayores recursos para gestionar la llegada de personas migrantes, concentraron numerosos comentarios hostiles. Entre ellos, destacan narrativas de “invasión” o sustitución demográfica, así como mensajes que cuestionan la asignación de recursos públicos a las políticas migratorias, presentando a las personas migrantes como beneficiarias indebidas de ayudas y contribuyendo a su estigmatización social.

El episodio prototípico de agresión sexual concentró durante el mes de julio el 5% del total de los contenidos analizados. Determinadas noticias relacionadas con agresiones sexuales, a raíz de sentencias judiciales, así como sucesos ocurridos durante fiestas locales, dieron lugar a un incremento de mensajes que asociaban de forma generalizada la violencia sexual con las personas de origen extranjero. Asimismo, persisten las narrativas que presentan a las personas migrantes como una amenaza para la seguridad de las mujeres, mediante expresiones como “sólo hacen salvajadas”, “las mujeres ya no podrán salir a la calle”, “los que entran vienen con ganas de violar y delinquir” o “sólo vienen a violar y a por paguitas”. Estas publicaciones contribuyen a reforzar estereotipos xenófobos y a legitimar discursos de exclusión hacia los grupos diana.

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)