



# BOLETÍN DE MONITORIZACIÓN DEL DISCURSO DE ODIO EN REDES SOCIALES

**2023**  
**1 ENERO**  
**28 FEBRERO**

El Observatorio Español del Racismo y la Xenofobia (OBERAXE) colabora desde 2017 con la Comisión Europea en los ejercicios de monitorización que se realizan en cumplimiento del Código de Conducta para la lucha contra la incitación ilegal al odio en Internet, que fue firmado con las empresas prestadoras de servicios de alojamiento de datos con mayor presencia en la Unión Europea.



Desde mayo de 2020, el OBERAXE monitoriza a diario el discurso de odio en España en cinco redes sociales (Twitter, Facebook, YouTube, Instagram y TikTok) y notifica aquellos contenidos considerados de odio racista y/o xenófobo, antisemita, antigitano e islamófobo o que pueden ser constitutivos de delito.

En el periodo comprendido entre el 1 de enero y el 28 de febrero de 2023, las plataformas han retirado el 38% de los contenidos notificados. Predomina la xenofobia hacia las personas originarias del norte de África (43%) y aumenta en 13 puntos porcentuales el discurso de odio islamófobo (23%).



# CONTENIDOS ANALIZADOS

## 695 contenidos de discurso de odio notificados

El número total de contenidos notificados este bimestre es de 695. En la Figura 1 se observa cierto predominio en el número de contenidos notificados a Facebook con 228 casos, seguido de Twitter (201), YouTube (104), Instagram (87), y TikTok (75). El desigual volumen de contenidos notificados obedece al distinto grado de dificultad para su identificación entre redes sociales.

## Las plataformas han retirado en total un 38% de los contenidos notificados

Las plataformas en su conjunto han retirado 262 contenidos de los 695 comunicados por el OBERAXE, es decir, el 38%. La respuesta de las plataformas resulta insuficiente. El 49% de los contenidos notificados infringen las normas de las propias plataformas.

## Las redes sociales han retirado un 21% de los contenidos notificados como *trusted flagger*\*

Del total de los 695 contenidos notificados, se han retirado 120 como usuario normal, lo que supone un 17%.

A su vez, mediante la vía *trusted flagger*, empleada este bimestre en 575 ocasiones,

las plataformas han retirado 142 publicaciones más, lo que supone un 21% adicional.

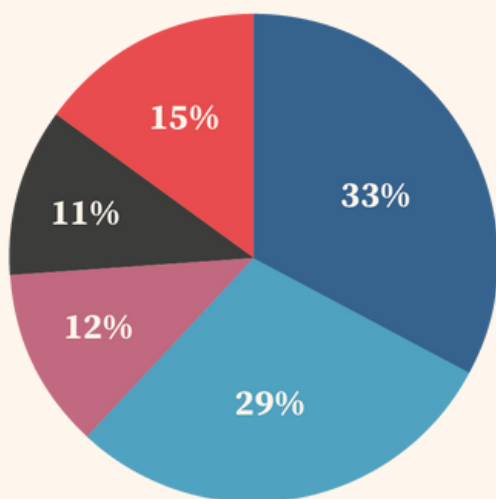
## Twitter sólo ha retirado el 9% del contenido notificado este bimestre

Al analizar los contenidos retirados por cada plataforma en relación con el total de casos notificados a cada una de ellas (Figura 2), TikTok es la red social con mayor tasa de retirada (93%). Le sigue Instagram, que ha retirado casi la mitad de los contenidos notificados (45%). Facebook ha retirado el 40%. YouTube ha bajado su tasa de retirada (41%) en 23 puntos porcentuales respecto al bimestre anterior (64%).

Twitter sólo ha retirado el 9% del contenido de discurso de odio racista y/o xenófobo comunicado, una tasa de retirada 46 puntos porcentuales menor respecto al bimestre anterior (55%).

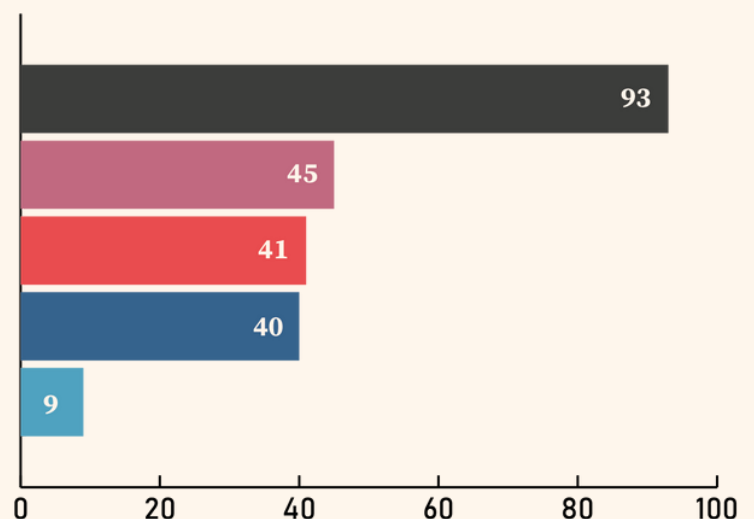
\*Conforme a la RECOMENDACIÓN (UE) 2018/334 DE LA COMISIÓN de 1 de marzo de 2018 sobre medidas para combatir eficazmente los contenidos ilícitos en línea se entiende como "comunicante fiable" (*trusted flagger*): aquella persona física o jurídica que un prestador de servicios de alojamiento de datos considere que posee competencias y responsabilidades particulares a efectos de la lucha contra los contenidos ilícitos en línea.

**FIGURA 1.**  
**CONTENIDOS COMUNICADOS**



● INSTAGRAM ● FACEBOOK ● TWITTER ● TIKTOK ● YOUTUBE

**FIGURA 2.**  
**% CONTENIDOS RETIRADOS**



# EVOLUCIÓN EN LA RETIRADA DE CONTENIDOS

La Figura 3 presenta los plazos en los que las plataformas retiran el contenido que se les ha notificado\*: 24 horas, 48 horas o una semana. Asimismo, incorpora el número de contenidos retirados tras la comunicación como *trusted flagger*.

## La mayor parte de la retirada de contenidos es a través de la figura del *trusted flagger*

Las plataformas continúan siendo más eficaces en la retirada de contenidos de discurso de odio cuando se emplea la vía de *trusted flagger*. El 54% del total de casos retirados (262) han sido efectuados a través de esta vía.

## TikTok es la red social más eficiente ante las notificaciones

TikTok ha retirado el 93% de los contenidos que se le han notificado en el periodo enero-febrero 2023 (75), siendo la plataforma que más contenidos ha retirado en relación a los notificados este bimestre. El 34% de los contenidos han sido eliminados en 24 horas, liderando también esta franja. El 43% fue retirado tras la notificación como *trusted flagger*.

## Facebook ha retirado el 40% del contenido notificado

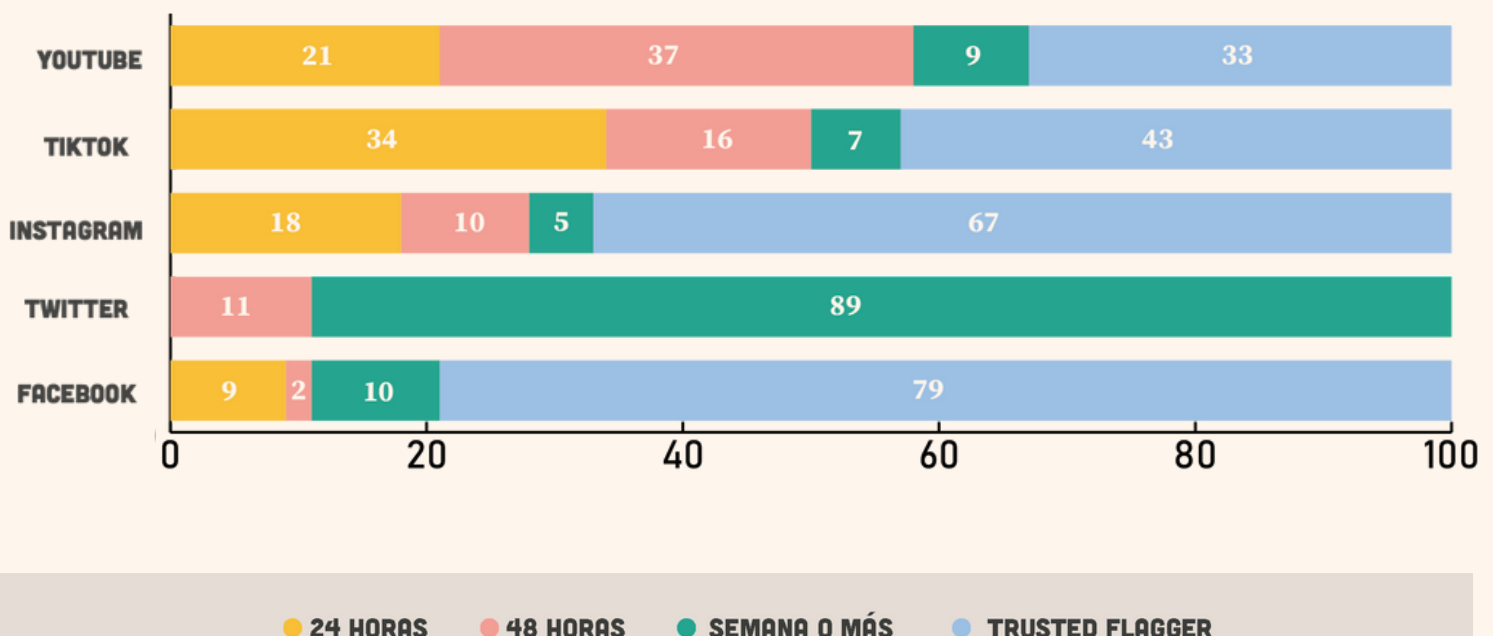
Facebook ha mejorado su tasa de retirada en 33 puntos porcentuales (40%) respecto al bimestre anterior (7%). Ha retirado el 79% del contenido notificado a la plataforma mediante la vía de *trusted flagger*, confirmándose como la red social más proactiva por esta vía en enero-febrero 2023. Instagram ha retirado el 67% de los casos notificados por la vía de *trusted flagger*. YouTube (33%), por su parte, se mantiene en cifras similares al bimestre anterior (30%).

## Twitter deja inoperativa la plataforma para la notificación de contenidos como *trusted flagger*

Desde la última semana del mes de diciembre del año 2022, el procedimiento de notificación de retirada a través de la vía *trusted flagger* de la plataforma Twitter no está operativo. No se han remitido contenidos por esta vía.

\*Si transcurrida una semana, la plataforma no retira el contenido denunciado como usuario normal, se procede a utilizar el procedimiento de comunicante fiable o *trusted flagger*.

**FIGURA 3.**  
**% TIEMPO DE RETIRADA DE CADA PLATAFORMA**



# CONTENIDO DE DISCURSO DE ODIOS

## La inseguridad ciudadana es el tema principal del discurso de odio

El 55% de los contenidos notificados a las plataformas hacen alusión a episodios de inseguridad ciudadana, reales o *fake news*. Este bimestre, se han dado otros acontecimientos puntuales (14%), relacionados con eventos deportivos, que han promovido discurso de odio hacia personas afrodescendientes o inmigrantes en general.

## El 55% de los contenidos deshumaniza o degrada gravemente a los grupos diana

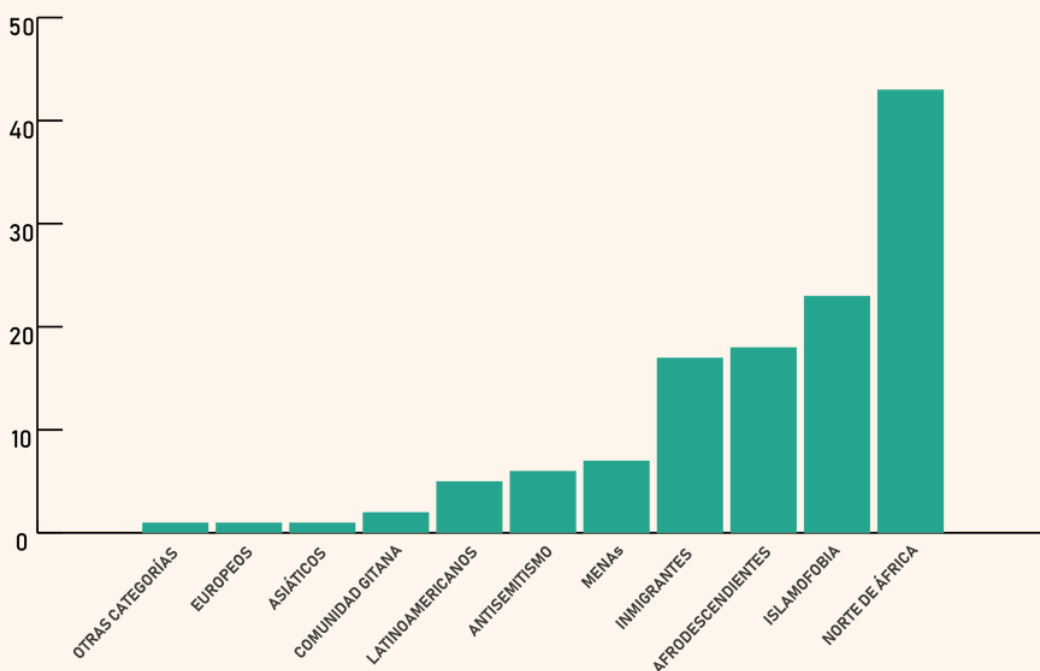
Aumentan en ocho puntos porcentuales respecto al bimestre noviembre-diciembre 2022 (47%) aquellos contenidos notificados que incluyen expresiones de deshumanización o degradación grave (55%). Un 43% del total de contenidos notificados promueve el descrédito a los grupos diana en base a atributos personales o del grupo.

## El 43% del discurso de odio notificado se dirige hacia personas originarias del norte de África

Se mantiene la hostilidad hacia las personas originarias del norte de África (43%), consolidando la tendencia de los dos boletines anteriores. El discurso islamófobo (23%) aumenta en 13 puntos porcentuales respecto a los datos del boletín anterior (10%). Este tipo de discurso de odio es sensible a incidentes concretos, como el ataque ocurrido el pasado 25 de enero de 2023 contra dos parroquias en el municipio de Algeciras.

Con un menor incremento porcentual (3 puntos) se localizan contenidos antisemitas (6%). Disminuye el discurso contra la comunidad gitana (2%) y el colectivo general de personas inmigrantes (17%). El discurso contra personas afrodescendientes (18%), latinoamericanas (5%) y los menores extranjeros no acompañados (7%) se mantiene estable respecto al boletín anterior.

**FIGURA 4.**  
**% COLECTIVOS VICTIMIZADOS EN EL DISCURSO DE ODIOS**



Edita y distribuye: Observatorio Español del Racismo y la Xenofobia.

NIPO: 121-23-006-3



GOBIERNO DE ESPAÑA

MINISTERIO DE INCLUSIÓN, SEGURIDAD SOCIAL Y MIGRACIONES

SECRETARÍA DE ESTADO DE MIGRACIONES

DIRECCIÓN GENERAL DE ATENCIÓN HUMANITARIA E INCLUSIÓN SOCIAL DE LA INMIGRACIÓN



Cofinanciado por la Unión Europea