

BOLETÍN DE MONITORIZACIÓN DEL DISCURSO DE ODIOS EN REDES SOCIALES

1 DE NOVIEMBRE - 31 DE DICIEMBRE (2021)



El **Observatorio Español del Racismo y la Xenofobia (OBERAXE)** realiza desde 2017 ejercicios de monitorización periódicos en cumplimiento del **Código de Conducta para la lucha contra la incitación ilegal al odio en Internet** que firmaron la Comisión Europea y algunas de las empresas prestadoras de servicios de alojamiento de datos con mayor presencia en todo el mundo, como **Twitter, Facebook, Youtube** o, más recientemente, **Instagram y TikTok**.

Además, desde **mayo de 2020** OBERAXE monitoriza diariamente el discurso de odio en España en estas redes sociales, comunicándoles aquellos contenidos considerados "de odio" o que pueden ser constitutivos de delito.

A continuación, se muestran los resultados obtenidos en la monitorización para el periodo comprendido entre el 1 de noviembre y el 31 de diciembre de 2021.

NÚMERO DE CONTENIDOS ANALIZADOS, COMUNICADOS Y RETIRADOS EN LAS REDES SOCIALES

El número de **contenidos** de discurso de odio identificados y comunicados a las plataformas Facebook, Twitter, Youtube, Instagram y TikTok durante este periodo ha sido de **660**, lo que supone una **disminución del 6,25%** con respecto al boletín bimensual anterior (1 septiembre- 31 de octubre 2021). De estos, **143 contenidos** fueron eliminados al ser notificados como usuario normal, por lo que la tasa de eliminación se ha situado en un **21,66%**, lo que supone un **descenso de 14,04 puntos porcentuales** respecto al periodo anterior.

En todas las redes sociales se ha utilizado, además, la vía de comunicación como **trusted flagger***. Este mecanismo se utiliza cuando ha transcurrido una semana desde las comunicaciones realizadas como **usuario normal**, y los contenidos no han sido eliminados. Tras las notificaciones como **trusted flagger**, las plataformas retiraron **133 contenidos adicionales**, un **20,15%** del total, por lo que la tasa de eliminación total asciende al **41,81%**, **cantidad inferior en 20,59 puntos porcentuales** respecto del periodo anterior.



**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.*

Por plataformas, de las **660 comunicaciones** realizadas, **191** corresponden a **Twitter** (29%), que es la red social a la que más comunicaciones se han realizado, seguida de **Youtube** con **184 comunicaciones** (28%) y **TikTok**, con **141** (21%). Asimismo, a **Facebook** se han remitido otras **92 comunicaciones** (14%) y **52 comunicaciones** (8%) a **Instagram**. La **Figura 1** refleja estos datos.

Figura 1. Porcentaje de comunicaciones realizadas a cada plataforma

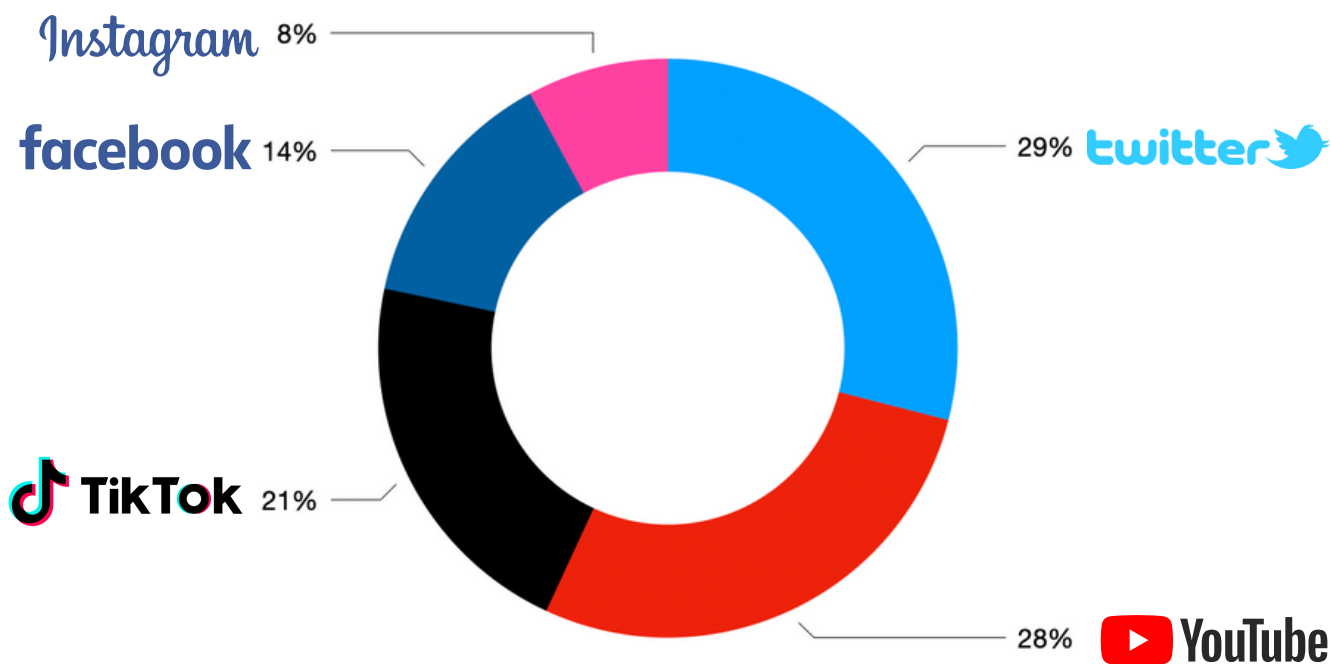
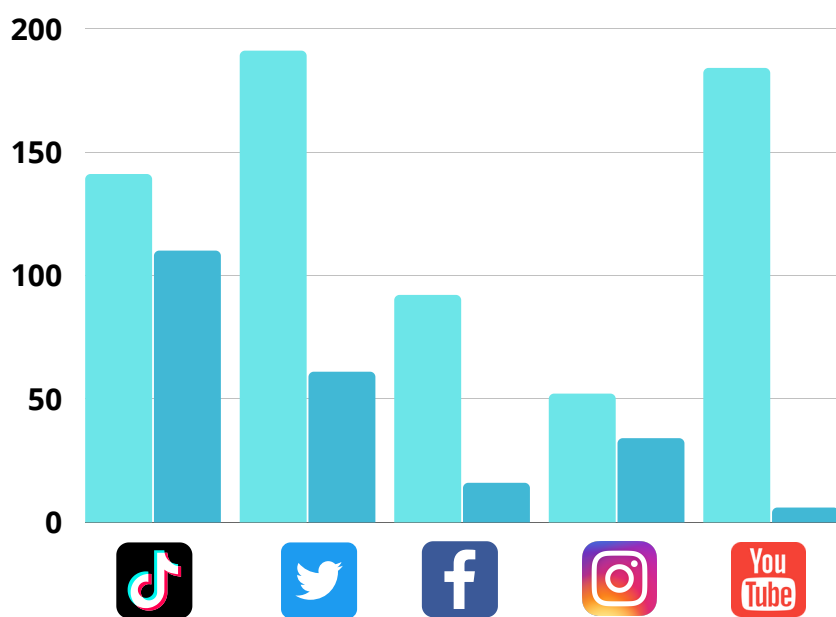


Figura 2. Número de contenidos retirados sobre el total de contenidos comunicados, según plataforma



En la Figura 2 se muestra el total de contenidos comunicados (en celeste) junto al número de contenidos retirados (en azul oscuro)



110 contenidos retirados de 141 comunicados (78,01%)



61 contenidos retirados de 191 comunicados (31,94%)



16 contenidos retirados de 92 comunicados (17,58%)



34 contenidos retirados de 52 comunicados (65,38%)



6 contenidos retirados de 184 comunicados (3,26%)

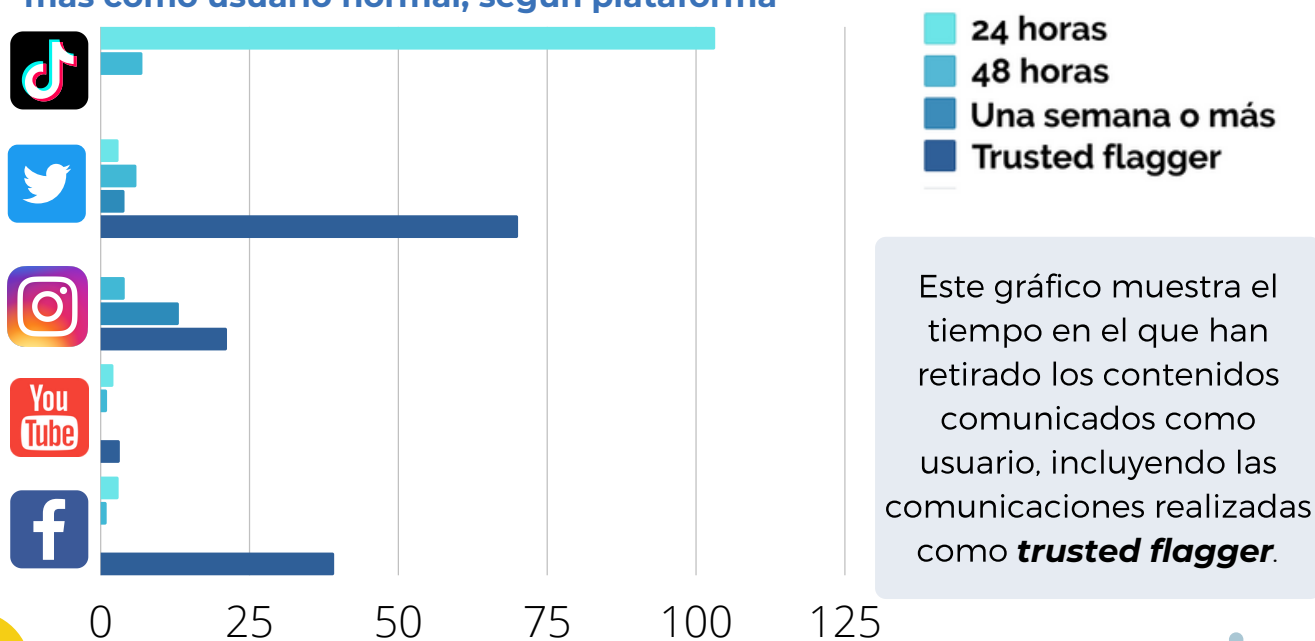
NÚMERO DE CONTENIDOS RETIRADOS A LAS 24H, 48H O UNA SEMANA Y COMO TRUSTED FLAGGER

En la monitorización, se realiza un seguimiento para comprobar cuánto tiempo les lleva a las plataformas tomar una decisión sobre la retirada o no del contenido comunicado. La Figura 3 recoge el **plazo temporal en el que las plataformas eliminan el contenido**: 24 horas, 48 horas, una semana o más. También se incorpora el número de comunicaciones eliminadas como **trusted flagger***.

Durante este ejercicio la eliminación de los contenidos ha sido más dispersa en el tiempo, y hay importantes cambios en algunas plataformas. No obstante, la tendencia a eliminar el grueso de los contenidos a las 24 horas desde su notificación permanece, a excepción de Instagram. Esta última, junto con Twitter, retiró la mayoría de contenidos tras comunicarlos como **trusted flagger**. Por plataformas, **TikTok** eliminó el **73,05%** de los contenidos comunicados en el primer día, seguido de **Facebook (3,26%)**, **Twitter (1,57%)**, **Youtube (1,09%)** e **Instagram (0%)**.

En el presente ejercicio destaca que plataformas como **Facebook y Twitter retiraron gran parte de los contenidos comunicados como trusted flagger** y no como usuario normal. Para la plataforma Twitter esta tendencia ha sido constante a lo largo del año (2021). En cambio para Facebook supone **un cambio en tendencia preocupante** dado que en los primeros meses del año gran parte del contenido se retiraba a las 24 horas de haber sido comunicado en calidad de usuario normal.

Figura 3. Contenidos retirados a las 24 horas, a las 48 horas, a la semana o más como usuario normal, según plataforma



*Si al cabo de una semana no se ha realizado la retirada por parte de la plataforma del contenido denunciado como usuario, se procede a utilizar la vía de comunicación cualificada o trusted flagger.



CONTENIDOS DE DISCURSO DE ODIOS SEGÚN MOTIVO DE DISCRIMINACIÓN EN CADA RED SOCIAL

La monitorización en este periodo muestra que Tik-Tok hace un esfuerzo importante por eliminar el contenido de odio de su plataforma, mientras que observamos un empeoramiento de la retirada por parte de Facebook. El **discurso de odio** está **muy influido por los flujos migratorios** que llegan a las fronteras, afectando sobre todo a la población de origen marroquí. Por otra parte, se mantiene un discurso contra los menores no acompañados basado en incidentes o **agresiones cuya autoría es desconocida** en la mayoría de las ocasiones, o que **no han sucedido en España**, pero que se atribuyen gratuitamente a menores no acompañados.

Los casos de **anti-inmigración** (Figura 4) han disminuido ligeramente con respecto al boletín anterior (33,7% en septiembre-octubre). Los contenidos contra los **menores extranjeros no acompañados** se mantienen como el tercer motivo de discriminación más común y también **descienden en 5,7 puntos porcentuales** respecto al periodo anterior (24,2% en septiembre-octubre).

Destaca el aumento **del contenido de carácter islamófobo**, que **se incrementa en 4,6 puntos porcentuales** (9,5% en el boletín anterior). Una cifra similar a la observada en el boletín de julio y agosto. Los contenidos de xenofobia hacia las personas de origen europeo (0,4% en septiembre-octubre) aumentan ligeramente, así como los contenidos antisemitas y de índole antigitana, que experimentan un ligero repunte con respecto al boletín bimensual anterior.

Figura 4. Contenidos de discurso de odio según motivo de discriminación

