

BOLETÍN MONITORIZACIÓN DE DISCURSO DE ODIO EN REDES SOCIALES



-1 DE OCTUBRE AL 31 DE DICIEMBRE 2020-



El Observatorio Español de Racismo y la Xenofobia (OBERAXE) participa desde 2017 en los ejercicios de monitorización del "Código de conducta para la lucha contra la incitación al odio en Internet" que firmaron la Comisión Europea y las principales empresas prestadoras de servicios de alojamiento de datos (**YouTube, Facebook, Twitter y Microsoft**) en 2016. Actualmente se han incorporado otras empresas.

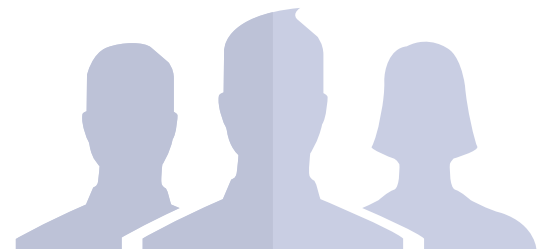
Desde mayo de 2020 el OBERAXE monitoriza diariamente en España el discurso de odio en Facebook, Twitter y YouTube, y notifica a las plataformas aquellos contenidos considerados "de odio". En este boletín trimestral se muestran los resultados obtenidos desde el 1 de octubre hasta el 31 de diciembre de 2020.

A partir del 1 de enero de 2021, la monitorización también se realizará en Instagram y TikTok.

NÚMERO DE CONTENIDOS ANALIZADOS, NOTIFICADOS Y RETIRADOS EN FACEBOOK, TWITTER Y YOUTUBE

Desde el 1 de octubre al 31 de diciembre se han notificado **424 contenidos** de discurso de odio identificados y analizados en Facebook, Twitter y YouTube. La mayoría de estos casos han sido denunciados como **usuario**. De ellos se han retirado **101**.

Tanto en Facebook, como en Twitter y YouTube se ha utilizado, además, la vía de denuncia de **trusted flagger**. Este mecanismo se utiliza una semana después de las notificaciones realizadas como **usuario**, siempre que los contenidos no son eliminados. Tras las notificaciones de **trusted flagger**, las plataformas retiraron **148 contenidos**.



**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.*

En el periodo analizado se han realizado 151 notificaciones a Facebook de las 424 totales, lo que supone el 35,6%. Le sigue muy de cerca Twitter con 150 contenidos notificados (35,3%), y por último, YouTube con 123 notificaciones (29,0%), como se aprecia en la Figura 1.

Figura 1. Porcentaje de notificaciones realizadas a cada plataforma

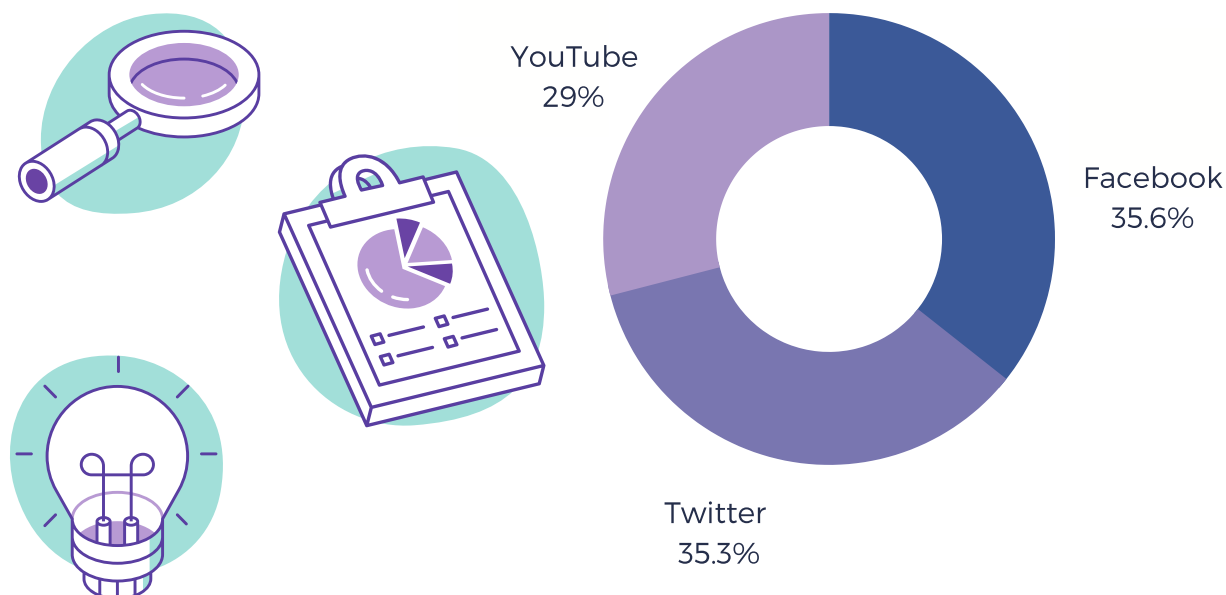
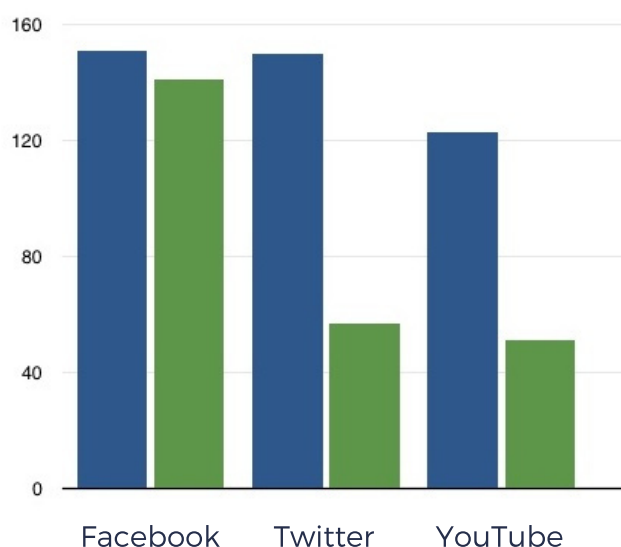


Figura 2. Número de contenidos retirados sobre el total de contenidos notificados, según plataforma



En la figura 2 se muestra el número de contenidos retirados (en verde) junto al total de contenidos notificados (en azul).

Facebook es la que más contenidos ha retirado, 141 sobre un total de 151 (93,4%). Mientras que Twitter solo ha retirado 57 contenidos sobre un total de 150 (41,5%) y YouTube 51 sobre 123 notificaciones (38,0%).



De los 151 contenidos notificados 141 fueron retirados



De los 150 contenidos notificados 57 fueron retirados



De los 123 contenidos notificados 51 fueron retirados

NÚMERO DE CONTENIDOS RETIRADOS A LAS 24H, 48H O 1 SEMANA

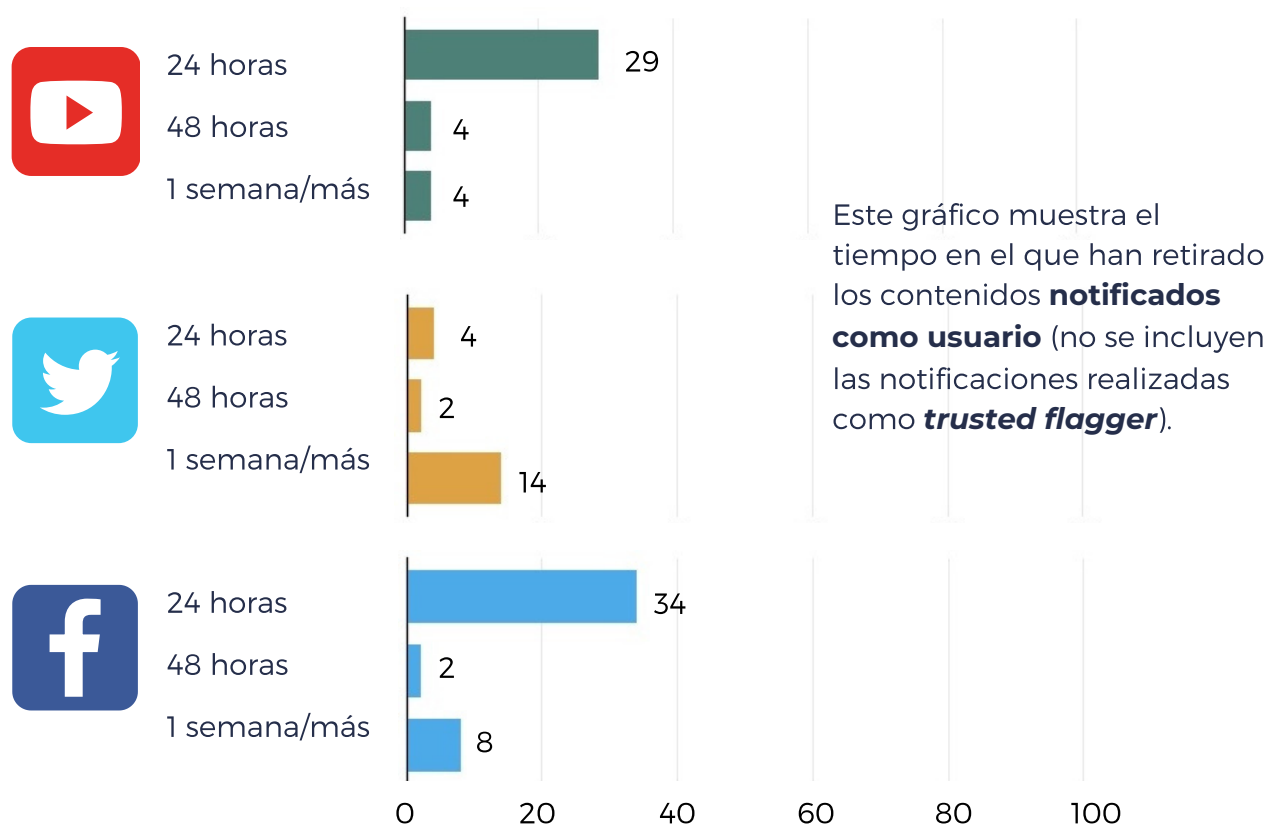
En el periodo de monitorización, se realiza un seguimiento para comprobar si el contenido denunciado se elimina a las 24 horas, 48 horas, a la semana o más (ver figura 3).

Facebook es la plataforma que, nuevamente, elimina más contenido en las primeras **24 horas**, con 34 retiradas, mientras que YouTube ha realizado 29 retiradas y Twitter 4.

Pasadas **48 horas**, YouTube eliminó 4 contenidos, por las 2 de Twitter y Facebook.

A la semana o más se han eliminado 14 contenidos en Twitter, 8 en Facebook y 4 en YouTube.

Figura 3. Contenidos retirados a las 24 horas, a las 48 horas, a la semana o más como usuario, 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 trusted flagger o comunicante fiable.*



CONTENIDOS DE DISCURSO DE ODO SEGÚN MOTIVO DE DISCRIMINACIÓN EN FACEBOOK, TWITTER Y YOUTUBE

Los resultados de la monitorización muestran que la agenda mediática sigue dirigiendo hacia dónde se inclina el discurso de odio en las plataformas estudiadas. En consonancia con los datos recopilados en el boletín anterior, la mayoría de las ofensas apuntan hacia inmigrantes en general (23,6%) y musulmanes en particular (19,1%).

En este trimestre, se ha detectado un aumento considerable del contenido discriminatorio dirigido hacia los asiáticos (15,8%, es decir, seis puntos más que en el periodo anterior) (ver figura 4).

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

