

BOLETÍN MONITORIZACIÓN DE DISCURSO DE ODIO EN REDES SOCIALES



-1 DE ENERO AL 28 DE FEBRERO 2021-

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 como TikTok.

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 bimensual se ha empezado a monitorizar también Instagram y TikTok. A continuación se muestran los resultados obtenidos desde el 1 de enero hasta el 28 de febrero de 2021.

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

Desde el 1 de enero al 28 de febrero se han notificado **399 contenidos** de discurso de odio identificados y analizados en Facebook, Twitter, YouTube, Instagram y TikTok. La mayoría de estos casos han sido denunciados como **usuario**. De ellos se han retirado **159**.

Tanto en Facebook, como en Twitter, YouTube e Instagram 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 **89 contenidos**. En total en este periodo se ha retirado el **62,2%** de los contenidos notificados.



**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 144 notificaciones a Facebook de las 399 totales, lo que supone el 36,1%. Le sigue Twitter con 101 contenidos notificados (25,3%), Instagram con 74 notificaciones (18,5%), YouTube con 52 notificaciones (13%), y por último, TikTok con 28 contenidos notificados (7%). Podemos apreciar estos datos en la Figura 1.

Figura 1. Porcentaje de notificaciones realizadas a cada plataforma

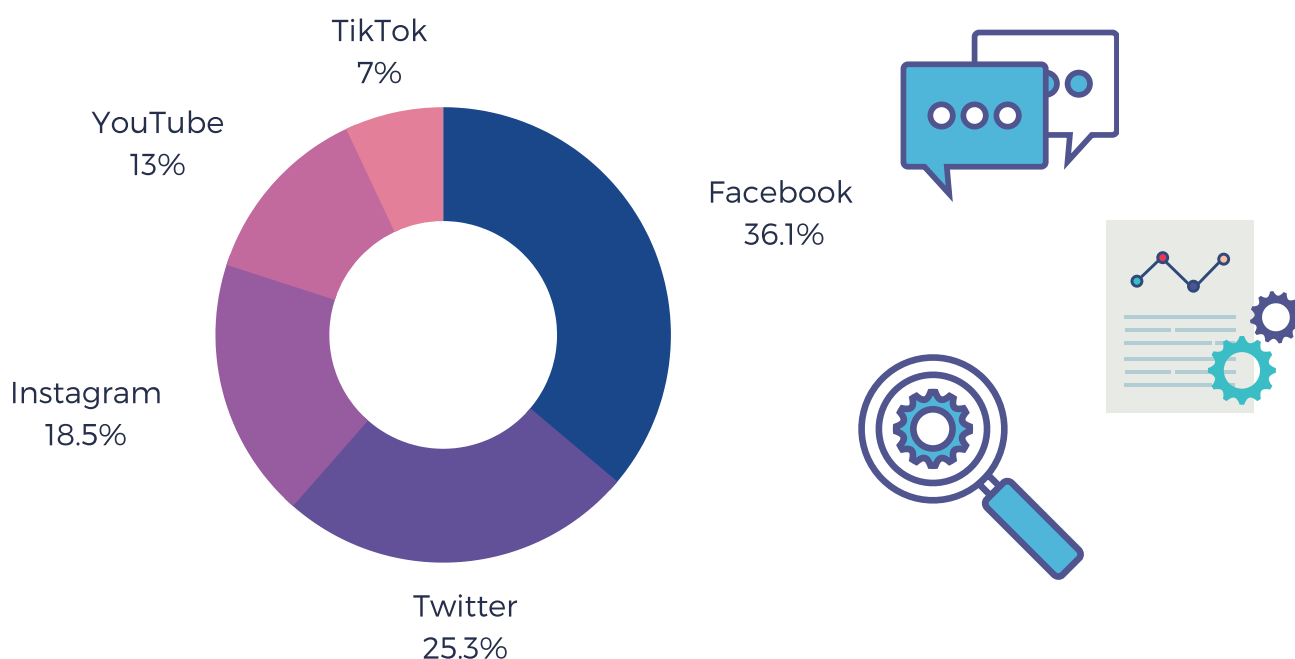
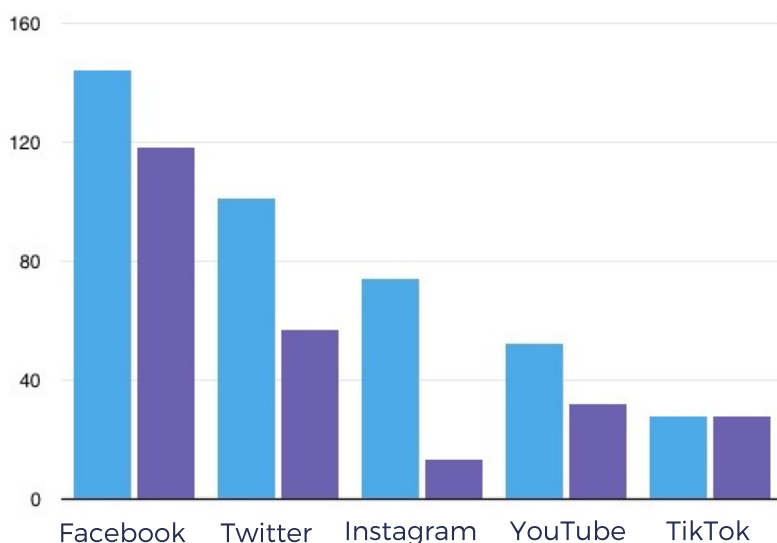


Figura 2. Número de contenidos retirados sobre el total de contenidos notificados, según plataforma



En la Figura 2 se muestra el total de contenidos notificados (en celeste) junto al número de contenidos retirados (en morado).



28 contenidos retirados de 28 notificados (100%)



118 contenidos retirados de 144 notificados (82%)



32 contenidos retirados de 52 notificados (61,5%)



57 contenidos retirados de 101 notificados (56,4%)



13 contenidos retirados de 74 notificados (17,6%)

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

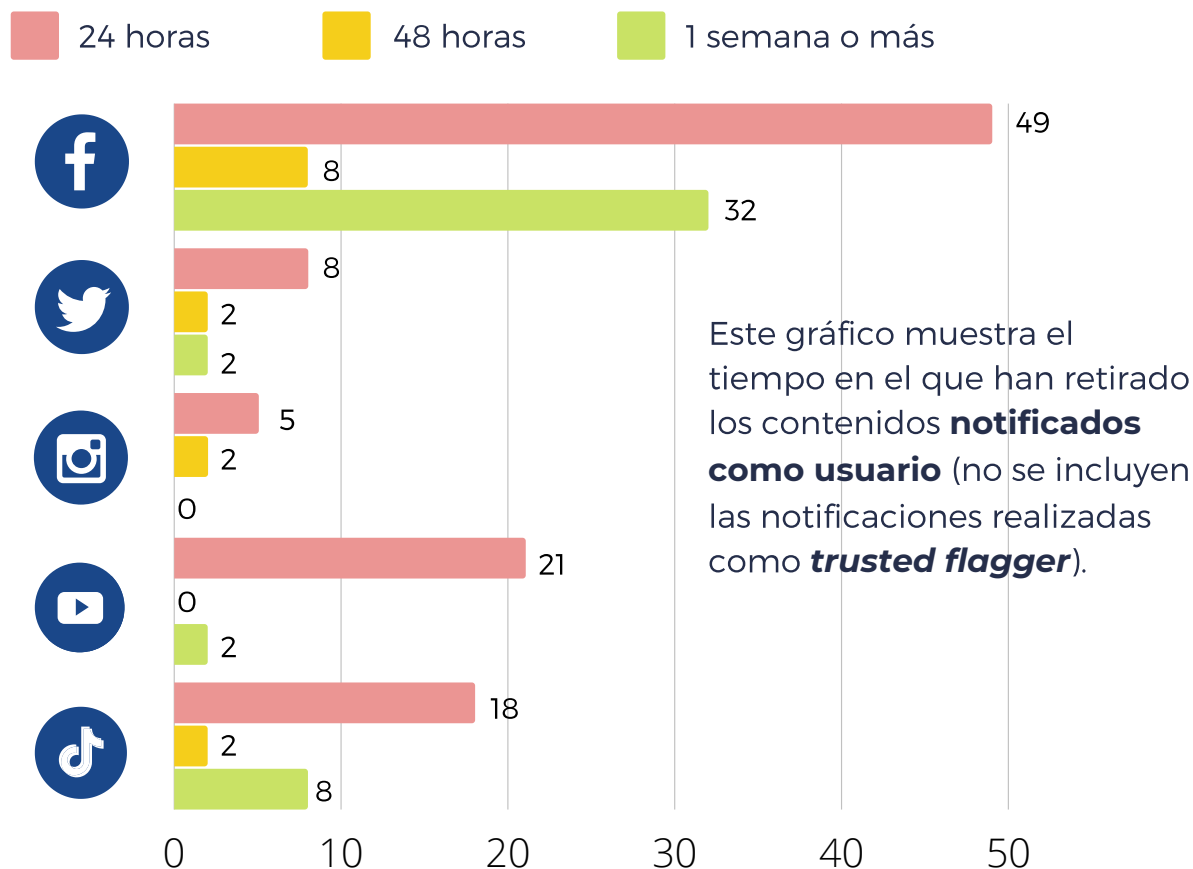
En el periodo de monitorización, se realiza un seguimiento para comprobar si el contenido denunciado se elimina a las 24 horas, 48 horas, a la semana o más (ver Figura 3).

Facebook elimina en las primeras **24 horas** 49 contenidos, mientras que YouTube ha realizado 21 retiradas, TikTok 18, Twitter 8 e Instagram 5 retiradas.

Pasadas **48 horas**, Facebook eliminó 8 contenidos mientras que TikTok, Twitter e Instagram eliminaron 2 contenidos. YouTube no eliminó contenido pasadas las 48 horas.

A la semana o más se han eliminado 32 contenidos en Facebook, 8 en TikTok y 2 en YouTube y Twitter. En Instagram no se eliminó contenido pasado este periodo de tiempo.

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, YOUTUBE, INSTAGRAM Y TIKTOK

Los resultados de la monitorización muestran que el discurso de odio en las plataformas estudiadas sigue la agenda mediática. En boletines anteriores se acentuó la xenofobia hacia asiáticos y hacia los menas mientras que en este la mayoría de las ofensas apuntan hacia inmigrantes (34,3%) en general y musulmanes (12,3%) en particular.

En el periodo de enero a febrero, se ha detectado un aumento considerable del contenido discriminatorio antisemita, sobre todo en Facebook. (10,8%, es decir, casi seis puntos más que en el periodo anterior) (ver Figura 4).

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

