

BOLETÍN MONITORIZACIÓN DE DISCURSO DE ODIOS EN REDES SOCIALES

-1 DE MARZO AL 31 DE ABRIL 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. En los últimos años se han incorporado otras empresas como TikTok.

Desde mayo de 2020 el OBERAXE monitoriza diariamente en España el discurso de odio en las redes sociales más utilizadas en el país, como Facebook, Twitter, YouTube, Instagram o Tiktok, y comunica a las plataformas aquellos contenidos considerados "de odio" en base a una serie de criterios objetivos. En el presente boletín bimensual se muestran los resultados obtenidos de la monitorización durante los meses de marzo y abril de 2021.

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

Desde el 1 de marzo al 31 de abril se han comunicado **455 contenidos** de discurso de odio identificados en Facebook, Twitter, YouTube, Instagram y TikTok, un 12,3% más de casos que en el último boletín bimensual. La mayoría de estos casos han sido comunicados como **usuario**, consiguiendo la retirada de **246** contenidos de odio, lo que supone un **54,1%** del total de contenidos detectados.

Tanto en Facebook, como en Twitter, YouTube e Instagram se ha utilizado, además, la vía de comunicación de **trusted flagger**. Este mecanismo se utiliza una semana después de las comunicaciones realizadas como **usuario** si los contenidos no han sido eliminados. Tras las comunicaciones de **trusted flagger**, las plataformas retiraron **113** contenidos más, por lo que en total en este periodo se han retirado **359** contenidos, un **78,9%** de los contenidos comunicados.



**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.*

De las **455** comunicaciones de contenido de odio, 120 comunicaciones corresponden a **Facebook**, lo que supone el **26,4%** del total. Le sigue **Twitter**, con 107 contenidos comunicados (**23,5%**), **TikTok** con 90 comunicaciones (**19,8%**), **Instagram** con 78 comunicaciones (**17,1%**), y por último, **Youtube** con 60 contenidos comunicados (**13,2%**). La Figura 1 recoge estos datos.

Figura 1. Porcentaje de comunicaciones realizadas a cada plataforma

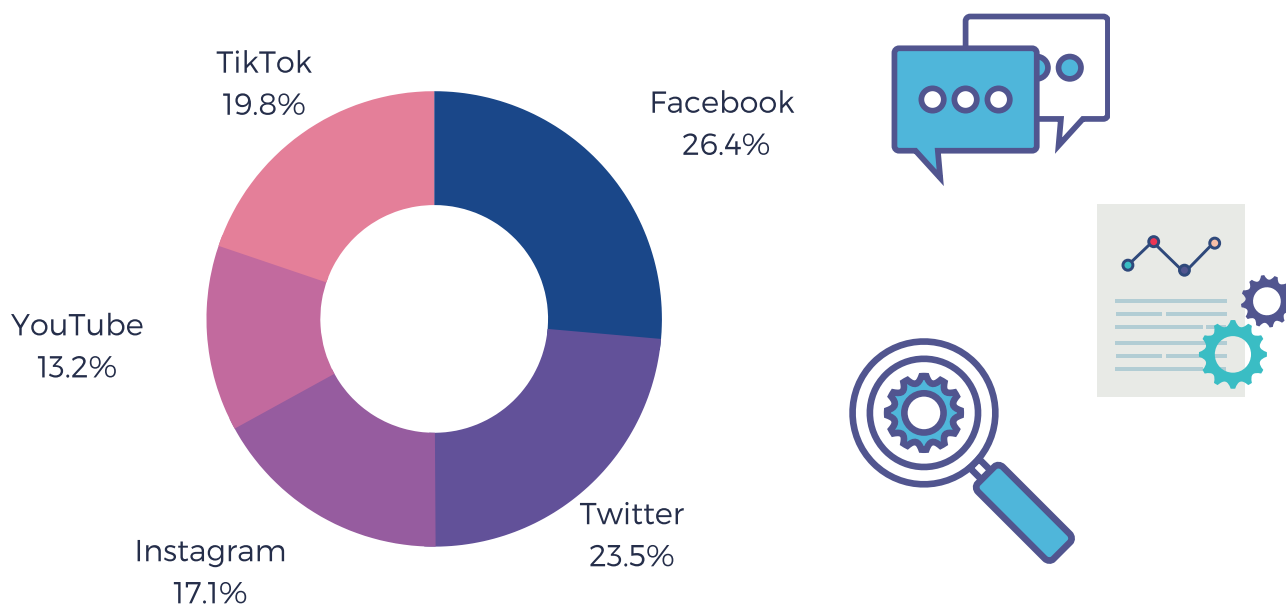
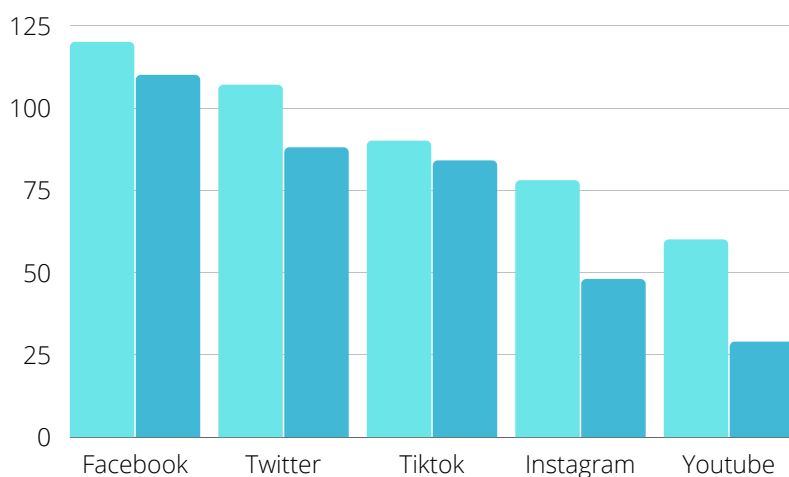


Figura 2. Número de contenidos retirados sobre el total de contenidos comunicados, según plataforma



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



84 contenidos retirados de 90 comunicados (93,3%)



110 contenidos retirados de 120 comunicados (91,7%)



88 contenidos retirados de 107 comunicados (82,2%)



48 contenidos retirados de 78 comunicados (61,5%)



29 contenidos retirados de 60 comunicados (48,3%)

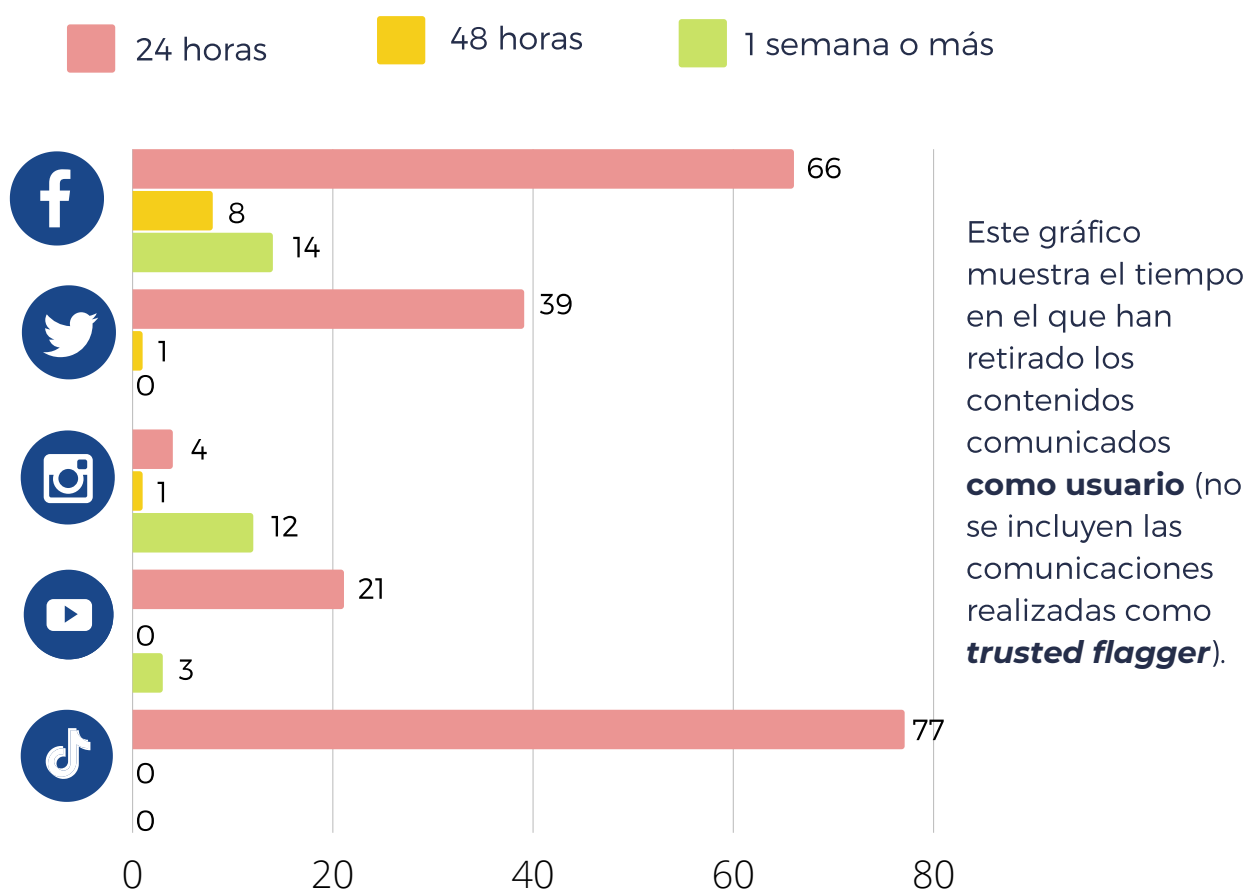
NÚMERO DE CONTENIDOS RETIRADOS A LAS 24H, 48H O 1 SEMANA

Como parte de la monitorización, se realiza un seguimiento para comprobar cuánto tiempo les lleva a las plataformas tomar una decisión sobre el contenido comunicado. La Figura 3 recoge el plazo temporal en el que las plataformas eliminan el contenido: 24 horas, 48 horas y/o una semana o más.

Facebook y Tiktok eliminaron en las primeras **24 horas** 77 contenidos cada una (64,2% y 85,6% respectivamente), mientras que Twitter realizó 39 retiradas (un 36%), Youtube 21 (35%) e Instagram 4 retiradas (5,1%).

Pasadas **48 horas**, Facebook eliminó 8 contenidos. Instagram y Twitter eliminaron 1 contenido adicional y Tiktok y Youtube no eliminaron nada. **A la semana o más** se eliminaron 14 contenidos en Facebook, 12 en Instagram y 3 en Youtube, mientras que Tiktok y Twitter no registraron novedad alguna.

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 es el reflejo de la agenda mediática y política. En este informe tenemos que destacar los casos de anti-inmigración (28,6%) y anti-inmigrantes menores (22%), blanco de los ataques que se han acentuado durante la reciente campaña electoral.

Destaca, asimismo, el porcentaje considerable de casos de islamofobia (11,4%) y xenofobia hacia los asiáticos (11%), que se ha detectado durante este ejercicio de monitorización (ver Figura 4).

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

