

BOLETÍN DE MONITORIZACIÓN DEL DISCURSO DE ODIOS EN REDES SOCIALES

1 DE MARZO - 30 DE ABRIL (2022)



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

Desde mayo de 2020 el **OBERAXE** monitoriza diariamente el discurso de odio en España en estas redes sociales, notificándoles 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 marzo y el 30 de abril 2022**.

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 **494**, lo que supone un **aumento del 27,0%** con respecto al boletín bimensual anterior (1 enero - 28 de febrero de 2022). De estos, **129 contenidos** fueron eliminados al ser notificados como usuario normal, por lo que la tasa de eliminación se ha situado en un **26,1%**, lo que supone un **incremento de 8,9 puntos porcentuales** respecto al periodo anterior.

En el ejercicio de monitorización 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 **116 contenidos adicionales**, por lo que la tasa de eliminación total asciende al **49,6%**, **valor superior en 27,7 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 **494 comunicaciones** realizadas, **203** corresponden a **Instagram** (41,1%), que es la red social a la que más reportes se han realizado, seguida de **Twitter** con **118** comunicaciones (23,9%) y **YouTube**, con **60** (12,1%). Asimismo, a **TikTok** se han remitido otras **58** comunicaciones (11,7%) y **55** a **Facebook** (11,1%). La Figura 1 refleja estos datos.

Figura 1. Porcentaje de comunicaciones realizadas a cada plataforma

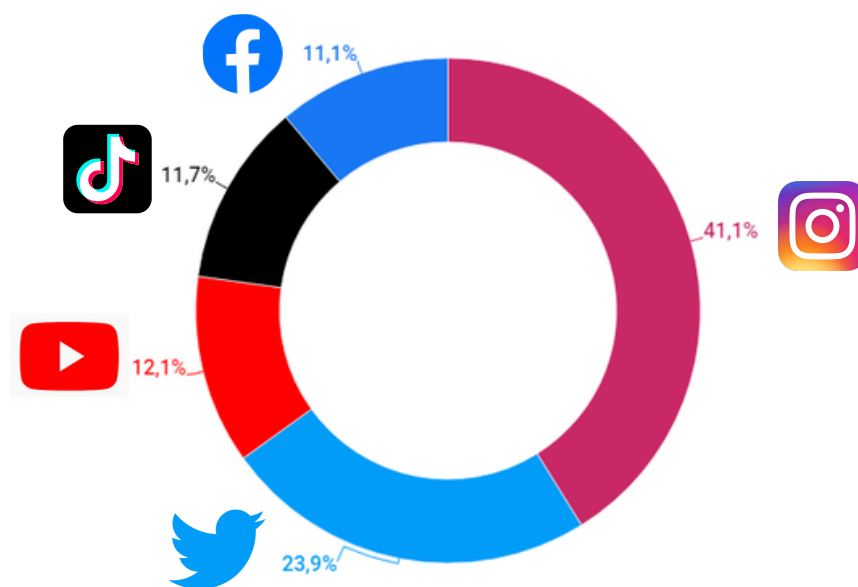
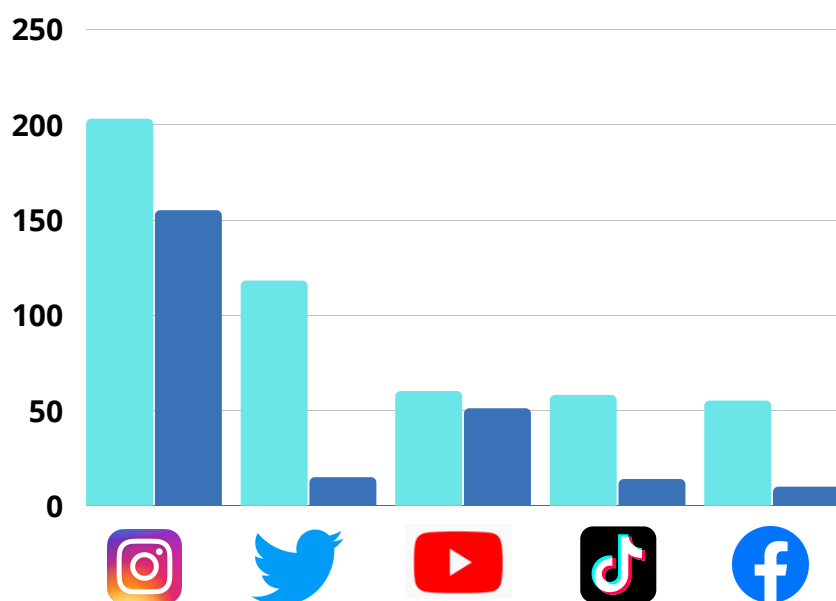


Figura 2. Número de contenidos retirados sobre el total de contenidos comunicados, según plataforma



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



155 contenidos retirados de 203 comunicados (76,3%)



15 contenidos retirados de 118 comunicados (12,7%)



51 contenidos retirados de 60 comunicados (85,0%)



14 contenidos retirados de 58 comunicados (24,1%)



10 contenidos retirados de 55 comunicados (18,2%)

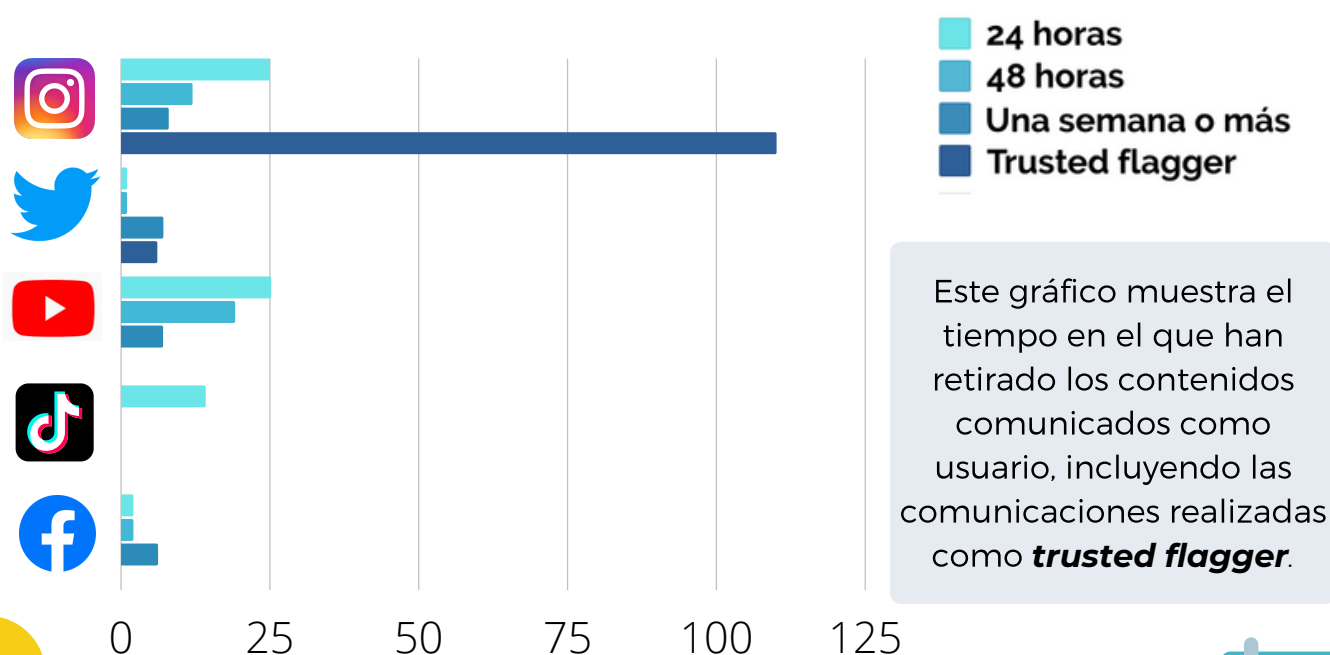
NÚMERO DE CONTENIDOS RETIRADOS A LAS 24H, 48H O UNA SEMANA Y COMO TRUSTED FLAGGER

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 contenidos varía notablemente en función de cada plataforma y no existe una clara tendencia generalizada. **YouTube** eliminó el **41,7%** de los contenidos comunicados en menos de 24 horas. En este plazo de tiempo destaca, en segundo lugar, **TikTok**, que eliminó en menos de un día el **24,1%** de los contenidos comunicados, lo que supone la totalidad de sus casos retirados. En dicho plazo, **Instagram** eliminó el **12,3%**; **Facebook**, el **3,6%**; y **Twitter**, el **0,9%**.

En el presente ejercicio **Instagram** ha mejorado su tasa de eliminación (**76,4%**), retirando gran parte de los contenidos como **trusted flagger** y no como usuario normal. **YouTube** registra una gran mejoría y ha retirado el **85,0%** de los contenidos comunicados frente a los cero casos retirados en el ejercicio anterior. **Twitter** y **Facebook** han eliminado un menor número de contenidos respecto al boletín bimensual anterior; y **TikTok** mantiene su tendencia.

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 las plataformas prestan **atención en muy diferente medida a la vía de comunicación de trusted flagger**. Solo **Instagram y Twitter** han retirado contenidos por este medio.

El **discurso de odio** sigue muy **influido** por los **flujos migratorios** que llegan a las fronteras, afectando sobre todo a la población de origen marroquí. Como se muestra en la Figura 4, los contenidos de **xenofobia** hacia las **personas de origen magrebí (31,0%) aumentan** 19,9 puntos porcentuales respecto a enero-febrero.

Los casos de **anti-inmigración** han **descendido** respecto al boletín anterior pasando de un 38,7% a un **30,4%** del total de contenidos reportados. Los casos de carácter **islamófobo** son el tercer motivo de discriminación más común (**13,4%**).

El **discurso de odio** contra los **menores no acompañados (6,9%) continúa reduciéndose** en relación con los boletines anteriores (10% en enero-febrero y 18,5% en noviembre-diciembre de 2021).

Los contenidos de **xenofobia** contra las **personas de origen europeo aumentan** 4,1 puntos porcentuales respecto al periodo anterior debido principalmente a la influencia de la invasión de Rusia a Ucrania. En Twitter se han monitorizado 7 casos aislados de **discurso de odio** contra la **población refugiada** procedente de **Ucrania**.

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

