



Discurso de odio

## **El Gobierno y las plataformas se comprometen a reunirse trimestralmente para analizar los discursos de odio en redes sociales y mejorar los datos de retirada**

- Representantes de Meta, YouTube y TikTok han participado junto a la ministra Elma Saiz, la secretaria de Estado de Migraciones y el nuevo director del OBERAXE en la primera reunión del grupo de trabajo
- Se han comprometido a reunirse cada tres meses y a organizar grupos de trabajo permanentes para cuestiones técnicas como definir mejor los discursos de odio y hacer un uso más eficiente de los canales prioritarios de las plataformas
- El informe anual del OBERAXE del 2024 revela que de los 2.870 contenidos notificados a las plataformas considerados como discurso de odio solo se retiró el 35%
- De enero a agosto de 2025, el OBERAXE ya ha detectado más de 600.000 contenidos de odio reportables, cifra considerablemente mayor que la del año pasado gracias al uso del sistema FARO – herramienta cedida al Ministerio a través de un convenio con LALIGA –
- Elma Saiz: “Recojo el guante para la celebración de una Cumbre contra el odio en las redes sociales, que analice y compare la situación de España con respecto a la de otros países”

**Madrid, 22 de septiembre de 2025.-** El Gobierno ha entregado a las plataformas digitales la memoria anual de monitorización del discurso de odio del año 2024 elaborada por el Observatorio Español del Racismo y la Xenofobia (OBERAXE), que en ese año notificó un total de 2.870 contenidos de carácter racista y solo el 35% fue retirado.



Precisamente la reunión celebrada este lunes, a la que han asistido representantes de Meta, YouTube y TikTok (X no ha podido asistir) ha concluido con el objetivo de **mejorar los datos y el tiempo de retirada de contenido de odio** mediante un trabajo colaborativo y coordinado entre todos los agentes con la celebración de **reuniones técnicas de seguimiento trimestrales** para analizar la evolución de los mensajes de contenido racista.

“Quiero agradecer a todas las plataformas digitales convocadas su asistencia a esta reunión. Hoy retomamos los trabajos del grupo de colaboración creado el pasado 30 de julio, **un espacio pionero en España y sin precedentes en la Unión Europea**. Frenar los bulos, la desinformación y el discurso de odio en las redes debe ser una tarea colectiva para que no se repitan imágenes como las de Torre Pacheco o las del ataque al centro de menores no acompañados de Monforte de Lemos”, ha señalado la **ministra Saiz**, que ha estado acompañada por la **secretaria de Estado Pilar Cancela** y el **nuevo director del OBERAXE, Tomás Fernández**.

**Las plataformas y el Ministerio se han comprometido a reunirse trimestralmente**, aunque se crearán grupos internos de trabajo para avanzar en cuestiones técnicas como las consecuencias de la implementación de la Ley de Servicios Digitales (DSA, por sus siglas en inglés) o encontrar una definición homogénea para el “discurso de odio”. Las plataformas han reiterado su interés por detener este tipo de mensajes y han compartido sus propios datos y herramientas para retirarlo antes de que llegue a visualizarse por los usuarios.

“Recojo el guante para la celebración de una Cumbre contra el odio en las redes sociales, que analice y compare la situación de España con respecto a la de otros países”, ha dicho la ministra Elma Saiz después de que las plataformas pusieran sobre la mesa la posibilidad de organizar un encuentro de alto nivel al respecto.



- **DATOS DEL INFORME ANUAL 2024**

Durante el pasado año, el Observatorio notificó a las cinco plataformas monitorizadas (Facebook, X, Instagram, TikTok y YouTube) **un total de 2.870 contenidos considerados como discurso de odio racista, xenófobo, antisemita, antigitano o islamófobo.**

Estos datos no han dejado de incrementarse durante 2025, en buena medida por la implementación de **un sistema de rastreo basado en inteligencia artificial, el Sistema FARO**, a través del cual OBERAXE ha detectado **614.833 contenidos reportables de enero a agosto**, con un pico registrado en el mes de julio que llegó a los 190.000.

Cabe recordar que los contenidos monitorizados y notificados por el sistema podrían ser constitutivos de delito, de infracción administrativa o que violan las normas de conducta de las plataformas digitales. No obstante, **los informes de OBERAXE señalan que el 65% de los contenidos reportados en 2024 y el 66% en la primera mitad de 2025 nunca fueron retirados de los canales de difusión de estas plataformas.**

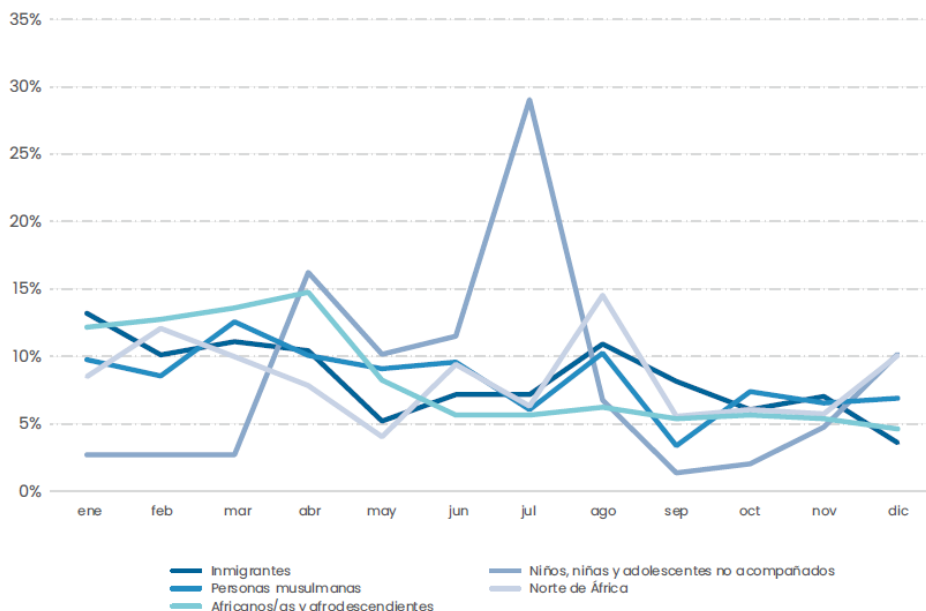
### **Personas originarias del norte de África, principal grupo diana**

Según los datos del Observatorio, en **2024 la hostilidad predominante hacia las personas originarias del norte de África se concentró en el 35% del total** de discursos de odio identificados, por detrás las personas africanas y afrodescendientes (24%) y la categoría genérica de personas inmigrantes (21%).

En referencia al grupo diana al que se dirige el discurso de odio, éste varía según la plataforma en la que este se difunde sin que se haya podido identificar ninguna relación causal. Así, vemos que en la plataforma X los discursos de odio principalmente van dirigidos a las personas del norte de África (30%) y en Facebook los discursos se dirigen con mayor frecuencia a inmigrantes en general (42%).



Gráfico 4. Evolución de los principales grupos diana en 2024



## Deshumanización de los grupos diana para incitar a la violencia

Entre las tipologías de discurso observadas en el informe de 2024, **la deshumanización se encuentra presente en el 37% de las comunicaciones notificadas** y en un 29% de los acontecimientos monitorizados se incita a la violencia con amenazas directas o indirectamente contra las personas migrantes o de origen extranjero. Además, **un 15% de los casos incitan a la expulsión del colectivo** y el 5% de los contenidos analizados elogian a quienes atacan al grupo diana, legitimando así la violencia y la discriminación.

El análisis también apunta en cuanto a la expresión del **discurso de odio agresivo explícito** es el más frecuente, observado en el 53% de los contenidos monitorizados.

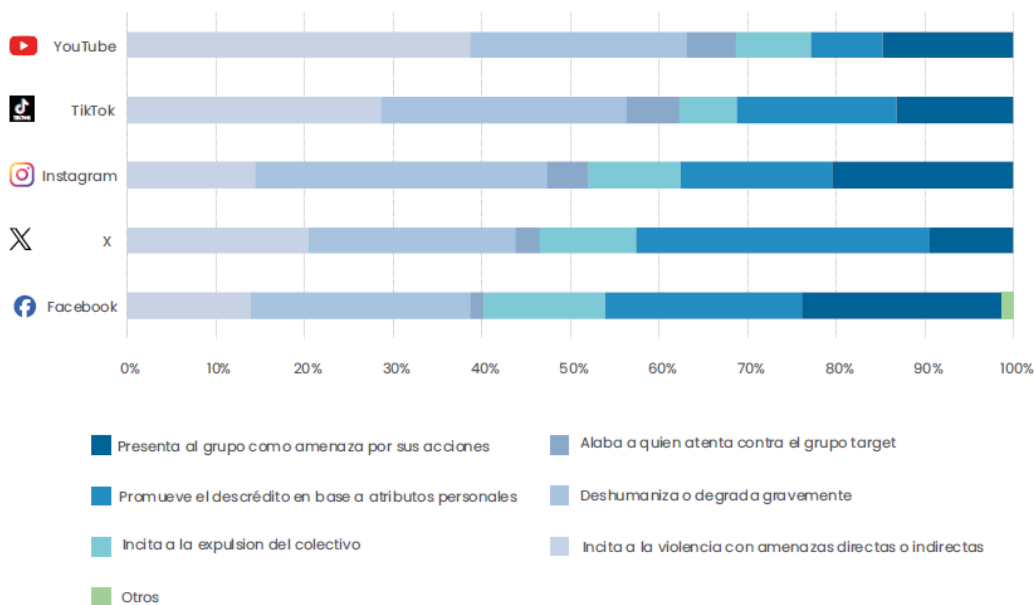
La tipología del discurso de odio también aporta diferencias entre las distintas redes sociales. En YouTube y TikTok predominan los contenidos que incitan a la violencia mediante amenazas directas o indirectas, con un 39% y un 29% respectivamente.

Por su parte, **los mensajes que deshumanizan a los grupos diana son más frecuentes en Instagram (33%) y Facebook (25%)**, donde también se observa una mayor presencia de contenidos que presentan al grupo objetivo como una amenaza



para la ciudadanía. En el caso de X, el tipo de discurso más recurrente es el que desacredita en función de atributos personales, con una prevalencia del 33%.

Gráfico 7. Distribución de la tipología del discurso de odio según red social



# Nota de prensa

Si en 2024 los contenidos que “deshumanizan o degradan gravemente” constituían un 37% de los contenidos notificados, en el primer semestre de 2025 la cifra se elevó hasta el 54%.

Más evidente aún es la tendencia de crecimiento en los datos que analizan el lenguaje expresado en este tipo de contenidos. Así, mientras que en 2024 el 53% de los contenidos monitorizados presentaba un lenguaje explícitamente agresivo, en el primer trimestre de 2025 la cifra se incrementó hasta un 70%.

El OBERAXE explica que el alto porcentaje de mensajes con lenguaje agresivo establece un patrón claro que evidencia la **consolidación de narrativas que**, al difundirse, **contribuyen a la estigmatización de las personas migrantes**, tanto en el entorno digital como en el público, reflejando su potencial impacto negativo.

## Retirada de contenidos del 4% en las primeras 24 horas



A lo largo de 2024, **las plataformas retiraron de sus canales tan solo el 35% de los contenidos que les fueron notificados**, en su mayor parte lo hicieron tras ser reportados mediante la figura de *trusted flagger* o comunicante acreditado como fiable. **De ese 35% de contenido retirado, solo el 4% se hizo en las primeras 24 horas, periodo crucial para evitar la viralización y minimizar el impacto de propio discurso.**

El Observatorio considera en su memoria que a pesar de las normas y mecanismos que establecen las plataformas en el marco del Código de Conducta y la normativa establecida por el Reglamento (UE) 2022/2065, conocido como Ley de Servicios Digitales (DSA), la retirada de contenido de discurso de odio sigue siendo insuficiente teniendo en cuenta que en un 96% de las comunicaciones se infringen las propias normas establecidas por cada una de las plataformas.

Tabla 1. Porcentaje de contenidos retirados según el tiempo transcurrido desde la notificación y según plataforma de internet, 2024

	Total contenido retirado %	Retirado a las 24 horas %	Retirado a las 48 horas %	Retirado a la semana %	Retirado <i>Trusted Flagger</i> %	No retirado %
Facebook	29%	1%	1%	5%	23%	71%
X/Twitter	14%	4%	1%	3%	7%	86%
Instagram	49%	2%	2%	7%	37%	51%
TikTok	69%	15%	3%	5%	47%	31%
YouTube	27%	0%	0%	1%	25%	73%
<b>Total</b>	<b>35%</b>	<b>4%</b>	<b>1%</b>	<b>4%</b>	<b>26%</b>	<b>65%</b>

## Discurso de odio: fenómeno estructural y persistente

Tras el análisis de los datos de 2024, el OBERAXE concluye que **la hostilidad en redes sociales hacia las personas de origen extranjero ha dejado de ser un fenómeno vinculado exclusivamente a hechos coyunturales**. Los datos del pasado año consolidan el **discurso de odio como una manifestación persistente, articulada en torno a estereotipos, bulos y polarización que encuentran en las redes sociales una vía rápida para su propagación**. Una proporción significativa del discurso de odio se mantiene activa incluso en ausencia de acontecimientos, lo que evidencia una base estructural de racismo y xenofobia.



En este contexto, se observan una serie de episodios prototípicos que actúan como desencadenantes recurrentes del discurso de odio, siendo el más predominante la vinculación entre inmigración e inseguridad ciudadana. No obstante, el Observatorio señala que en torno al 40% de los contenidos referentes a la inseguridad ciudadana son solo una percepción y no están basados en hechos verídicos, actuales y ocurridos en España, sino que se refieren a información falsa, bulos o incidentes descontextualizados. **La catástrofe de la DANA** ocurrida el pasado 29 de octubre es un ejemplo en el que **el discurso de odio focalizó la criminalización de personas del norte de África, musulmanes y comunidad gitana, relacionándolos con saqueos y robos.**

Por otro lado, las políticas públicas continúan siendo un blanco frecuente de discursos de odio, que aprovechan decisiones administrativas para alimentar teorías de privilegio institucional hacia los migrantes.

También el **ámbito deportivo** ha sido otro escenario donde se ha expresado de manera intensa el racismo y la xenofobia en redes sociales durante el 2024. El 4% de los contenidos monitorizados en 2024 estuvo dirigido principalmente a futbolistas, como en el caso de Vinícius Júnior, cuyo episodio de discriminación volvió a poner en el centro del debate el racismo tanto en los estadios como en las plataformas digitales.

### **Herramienta cedida por LALIGA en virtud de un convenio suscrito en 2024**

El **Sistema FARO** es una innovadora metodología de monitorización diseñada para la identificación y el análisis a tiempo real de los contenidos de discurso de odio con motivaciones racistas, xenófobas, islamófobas, antisemitas y antigitanas en las principales redes sociales. FARO, cuyo origen es una herramienta cedida por LALIGA en virtud del convenio de colaboración suscrito en octubre de 2024, representa un importante paso adelante en la lucha contra la intolerancia en el entorno de las redes sociales.

Datos de FARO en tiempo real:

<https://www.inclusion.gob.es/web/oberaxe/discurso-de-odio>