

1 ENERO - 31 MARZO 2025

Boletín de monitorización del discurso de odio en redes sociales



La monitorización del discurso de odio realizada por el Observatorio Español del Racismo y la Xenofobia (OBERAXE) desde el año 2020 consiste en la identificación, análisis y notificación a las plataformas de contenidos de discurso de odio con motivación racista, xenófoba, islamófoba, antisemita y antigitana, publicados en cinco plataformas de redes sociales (Facebook, Instagram, TikTok, YouTube y X); y que puedan ser constitutivos de delito, de infracción administrativa, o que infrinjan las normas de uso de las propias plataformas de prestación de servicios digitales.

La base de la metodología de monitorización parte del modelo establecido en los ejercicios de evaluación del cumplimiento del *Código de Conducta para la lucha contra la incitación ilegal al odio en Internet*, firmado en 2016 por la Comisión Europea junto con las plataformas de prestación de servicios digitales; y renovado en 2025 a través del *Código de Conducta +*.

El convenio de colaboración firmado entre el Ministerio de Inclusión, Seguridad Social y Migraciones y LALIGA, ha permitido al OBERAXE profundizar y multiplicar el alcance del trabajo realizado gracias al Sistema FARO (Filtrado y Análisis de Odio en las Redes Sociales). Un sistema que aplica la inteligencia artificial, entrenada en el Monitor para la Observación del Odio en el Deporte (MOOD) de LALIGA, a la metodología, especialización y experiencia acumulada por el OBERAXE. El Sistema FARO permite identificar y analizar en tiempo real los discursos de odio racistas y xenófobos en redes sociales, facilitando así la detección de los acontecimientos sociopolíticos que suscitan y amplifican estos discursos.

Nota: Los datos presentados en este boletín deben ser interpretados con cautela, ya que corresponden a los primeros resultados del Sistema Faro que se lanzó en marzo de 2025, y, por tanto, todavía está en fase de optimización de la herramienta de inteligencia artificial.

Contenidos monitorizados

En el periodo entre el 1 de enero y el 31 marzo de 2025, el monitor FARO ha detectado 171.273 contenidos de discurso de odio reportables de los cuales se han realizado 551 notificaciones a las plataformas, y estas han retirado el 34% de los contenidos reportados. El discurso de odio se ha dirigido principalmente a las personas del norte de África y a las musulmanas. En el primer trimestre del año, diversos acontecimientos han suscitado discursos de odio, siendo el Ramadán uno de los episodios más representativos.

171.273

Mensajes detectados

551

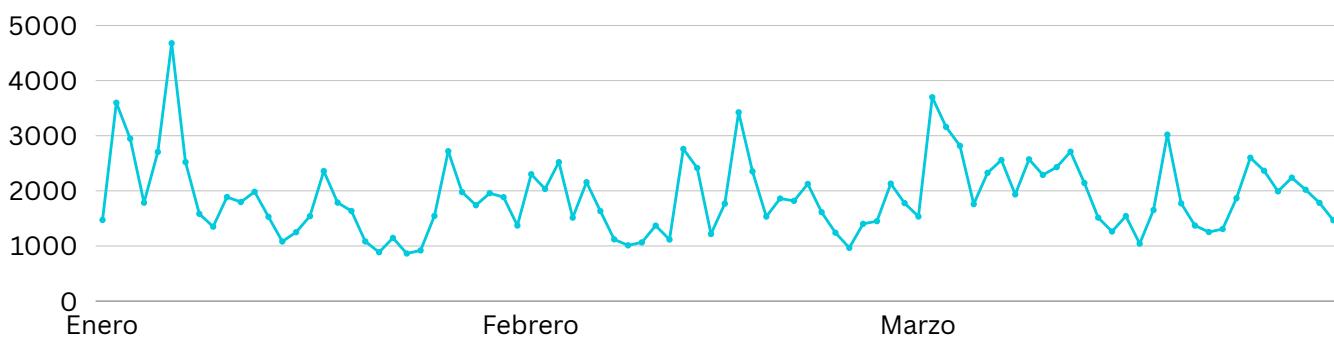
Mensajes reportados

34%

% Mensajes retirados

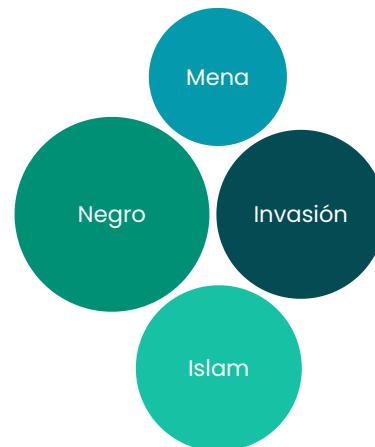
Evolución de los contenidos detectados

En el análisis del primer trimestre de 2025, se observa una fluctuación significativa en la detección de mensajes con contenido de odio. Enero presenta una disminución progresiva tras un máximo inicial, mientras que febrero experimenta un repunte significativo, con picos de hasta 1.019 mensajes diarios. Marzo mantiene una elevada incidencia, alcanzando su punto máximo el día 6 con 1.151 registros. Estas variaciones sugieren una posible relación causal con eventos sociales que se detallan en el apartado de episodios prototípicos.



Palabras clave

El monitor FARO ha detectado que las principales palabras clave en los contenidos que contienen discurso de odio durante este trimestre son las siguientes:

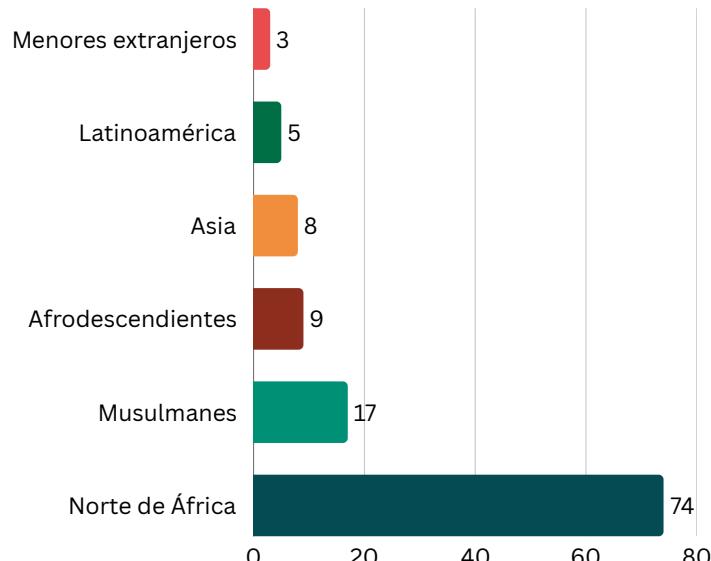


Características del discurso de odio

Grupo diana

Las personas de origen del norte de África siguen constituyendo el colectivo más atacado según los contenidos de discurso de odio reportados (74%). A este grupo le siguen, las personas musulmanas (17%) y las personas africanas y afrodescendientes (9%).

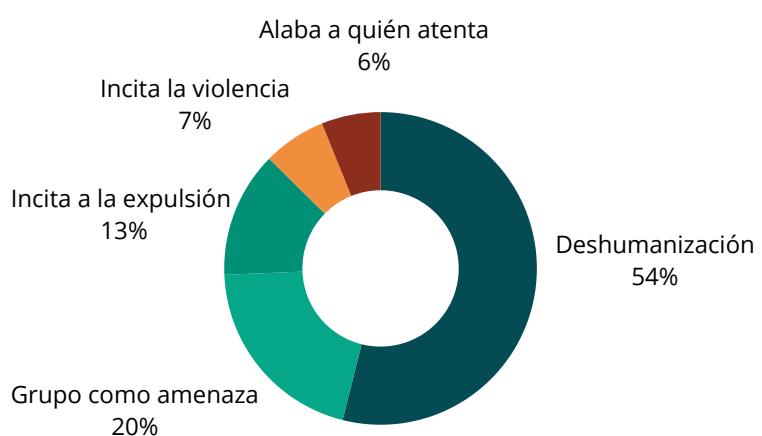
El discurso de odio dirigido hacia los menores extranjeros en este trimestre ha sido del 3%. Los contenidos vinculan directamente a los niños, niñas y adolescentes no acompañados con la inseguridad ciudadana.



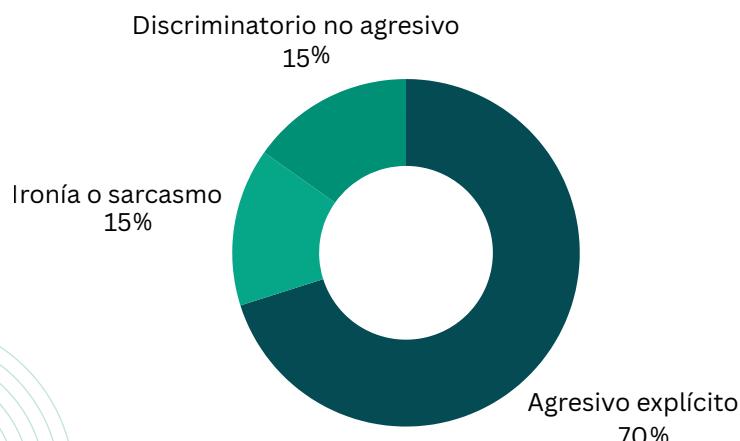
8 % El incremento del discurso de odio hacia las personas de origen asiático podría estar ocasionado por el conflicto comercial entre EEUU y China

Tipo de contenido

El 54% de los contenidos se caracterizan por la deshumanización o degradación de los diferentes grupos diana, a través de mensajes despectivos basados en estereotipos. Además, el 20% los presentan como una amenaza para la ciudadanía, y se incita a la expulsión de las personas de origen extranjero en un 13%. Un 6% de los contenidos analizados incita explícitamente a la violencia contra estas personas, mientras que un 6% elogia a quienes promueven estos discursos.



Expresión del lenguaje

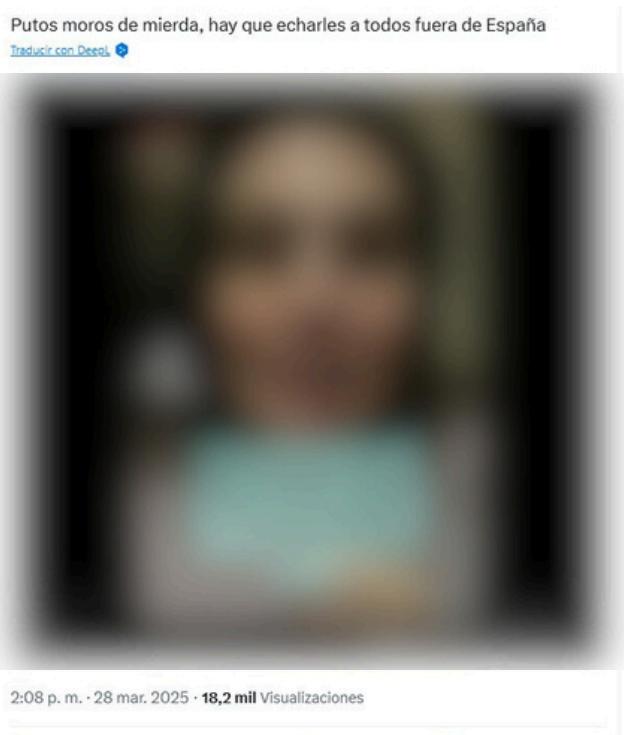


En el primer trimestre de 2025, el 70% de los contenidos presentan un lenguaje explícitamente agresivo. Le siguen, los contenidos que emplean la ironía o el sarcasmo (15%) y lenguaje discriminatorio no agresivo (15%).

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.

 94 % La mayoría de los contenidos de discurso de odio van dirigidos a hombres.

Contenido viral



En esta publicación de la plataforma X que ha tenido 18.2 mil visualizaciones, se observa el discurso de odio contra personas del norte de África/musulmanas al usar las palabras “*putos moros de mierda*”.

Además, en el video que procede de un contenido de la plataforma TikTok se utiliza un tono agresivo explícito, para incentivar a la expulsión del colectivo con palabras como “*hay que echarles a todos fuera de España*”.

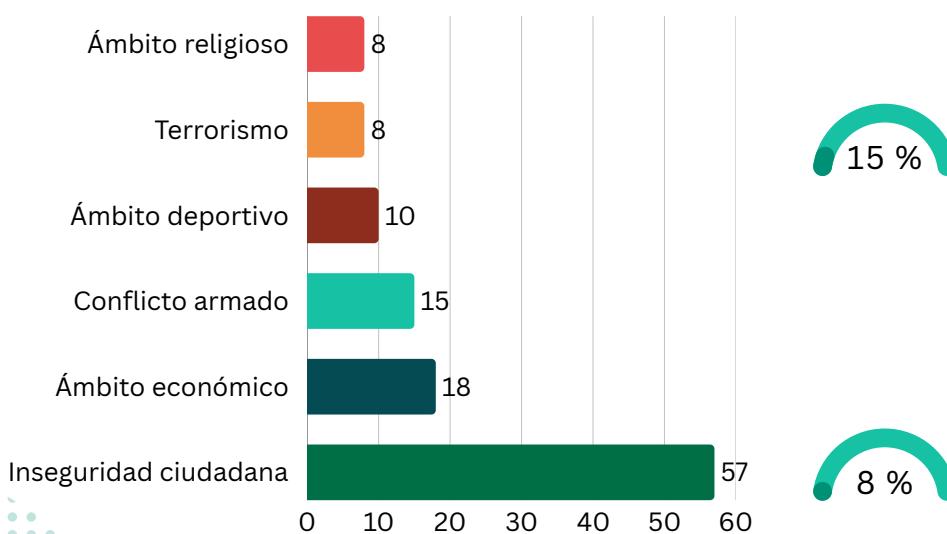
Episodios que suscitan discurso de odio

En el primer trimestre de 2025, se detectaron varios episodios prototípicos de discursos de odio en redes sociales relacionados principalmente con la inseguridad ciudadana, las políticas públicas y actos religiosos.

Los contenidos monitorizados revelan que el 57% de los discursos de odio estaban vinculados a la percepción de inseguridad ciudadana (robo/violencia), siendo dirigidos mayoritariamente hacia personas del norte de África y musulmanes. Algunos casos de inseguridad ciudadana monitorizados han estado relacionados con el atropello en Múnich (Alemania), el pasado 13 de febrero, que generó reacciones políticas polarizadas que vinculan la inmigración con la inseguridad en Europa.

Además, se ha evidenciado una creciente relación entre los discursos de odio y las políticas públicas de ámbito económico, representando el 18% de los casos monitorizados. Surgieron comentarios hostiles en torno a las políticas de acogida de reparto de menores extranjeros en las comunidades autónomas, a través de los cuales se cuestionaba la edad de los menores y se les vinculaba con la inseguridad, además se promovió una narrativa de rechazo y exclusión. Así como a la llegada de embarcaciones a las costas.

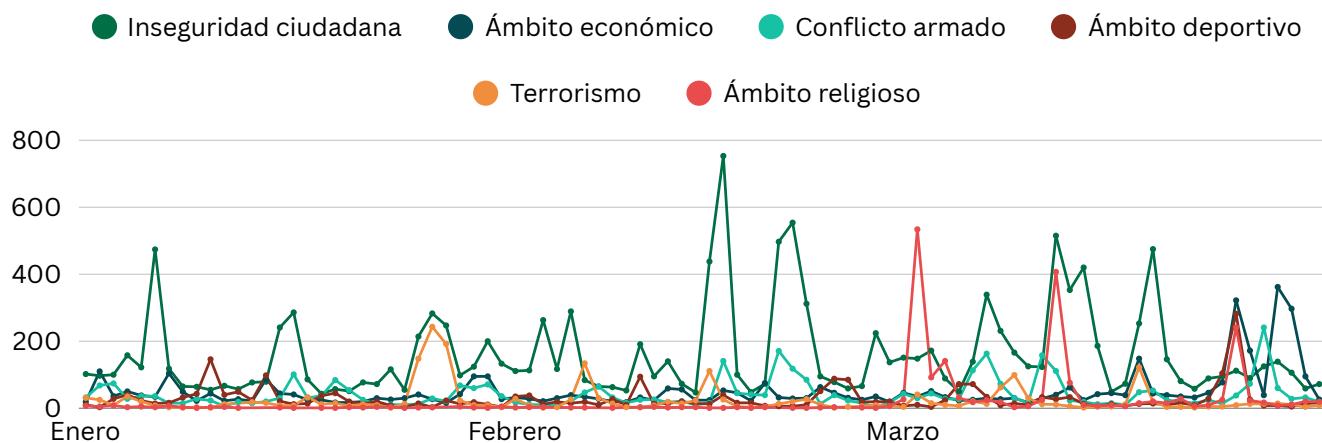
En el ámbito religioso, que representó el 8% de los contenidos, se observó un patrón creciente de comentarios que vinculan la religión islámica con una amenaza para la cohesión social, especialmente en torno a la celebración del Ramadán. Esta práctica religiosa ha sido utilizada para cuestionar la integración de las personas musulmanas en España. Como resultado, ha emergido un discurso agresivo que incita a la expulsión de este grupo, presentándolo como una amenaza para la ciudadanía.



La mediación de EE. UU. en la guerra Rusia-Ucrania podría estar vinculada al aumento del discurso de odio relacionado con el conflicto armado

Contenidos que utilizan el Ramadán para fomentar el discurso de odio hacia las personas musulmanas.

Evolución de los episodios prototípicos



En el primer trimestre de 2025, diversos sucesos sociopolíticos han suscitado el discurso de odio en redes sociales. El gráfico muestra los picos más significativos vinculados a episodios que han generado mayor contenido de odio. Estos picos reflejan cómo ciertos eventos son utilizados para alimentar estereotipos y promover la hostilidad hacia los diferentes grupos diana.

A lo largo del trimestre, la inseguridad ciudadana se ha utilizado como un pretexto para difundir discurso de odio. Un pico relevante ocurrió el 28 de enero, relacionado con la conversión de una mujer española al islam, que desató una serie de comentarios que asociaban la conversión con el "plan Kalergi" y el yihadismo. Por otro lado, el asesinato de Salwa Momika en Suecia el 30 de enero generó comentarios agresivos contra la religión islámica y la inmigración. Asimismo, un suceso ocurrido en Villach (Austria) fue utilizado para reforzar estereotipos negativos sobre las personas del norte de África.

Del mismo modo, el desalojo de un imán en Salt (Girona) el 13 de marzo fue vinculado con la inseguridad ciudadana, generando desconfianza hacia las personas musulmanas, este hecho alimentó estereotipos sobre radicalización y falta de integración creando un clima de estigmatización y hostilidad.

Por otro lado, el Ramadán se ha convertido en un foco de estigmatización, como se evidenció en marzo con los ataques a los futbolistas Brahim Díaz y Lamine Yamal, cuyo origen musulmán y práctica del Ramadán fueron utilizados como pretexto para incitar al odio.

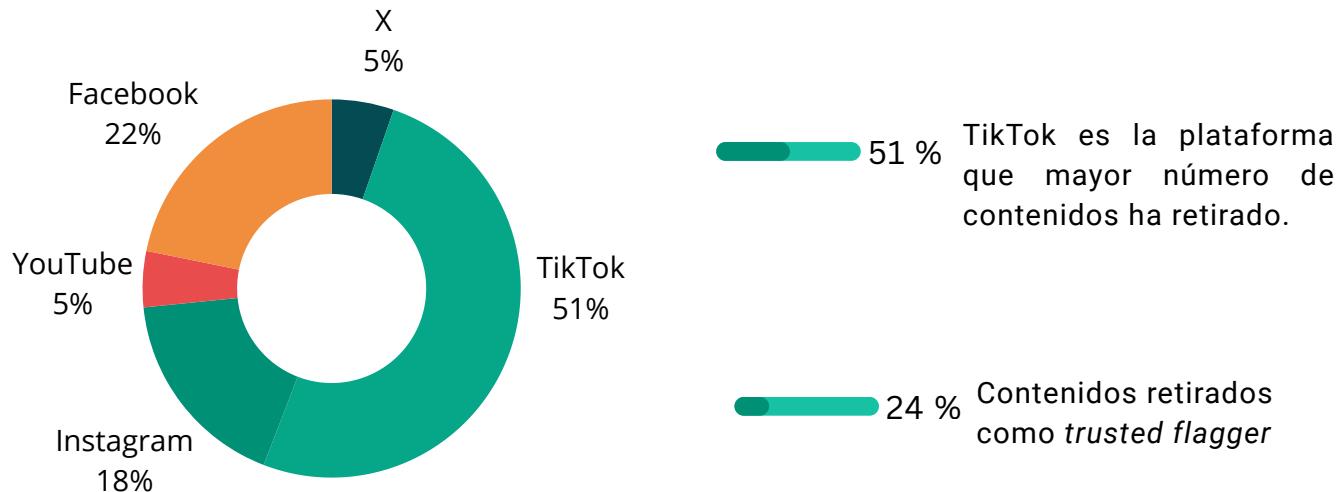
También, en el ámbito de las políticas públicas el reparto de fondos para la inmigración y el de niñas, niños y adolescentes no acompañados ha generado discursos que asociaban la inmigración con la inseguridad y la radicalización, promoviendo la expulsión de las personas de origen extranjero de la sociedad.

Reacciones de las plataformas

Las plataformas en su conjunto han retirado 188 contenidos del total de las 551 notificaciones realizadas por el OBERAXE, es decir, el 34%. Por un lado, se han retirado 39 notificaciones como usuario normal (10%). Las plataformas han retirado un 3% de las notificaciones en un plazo inferior a las 24 horas, un 3% en las 48 horas y un 4% el transcurso de una semana.

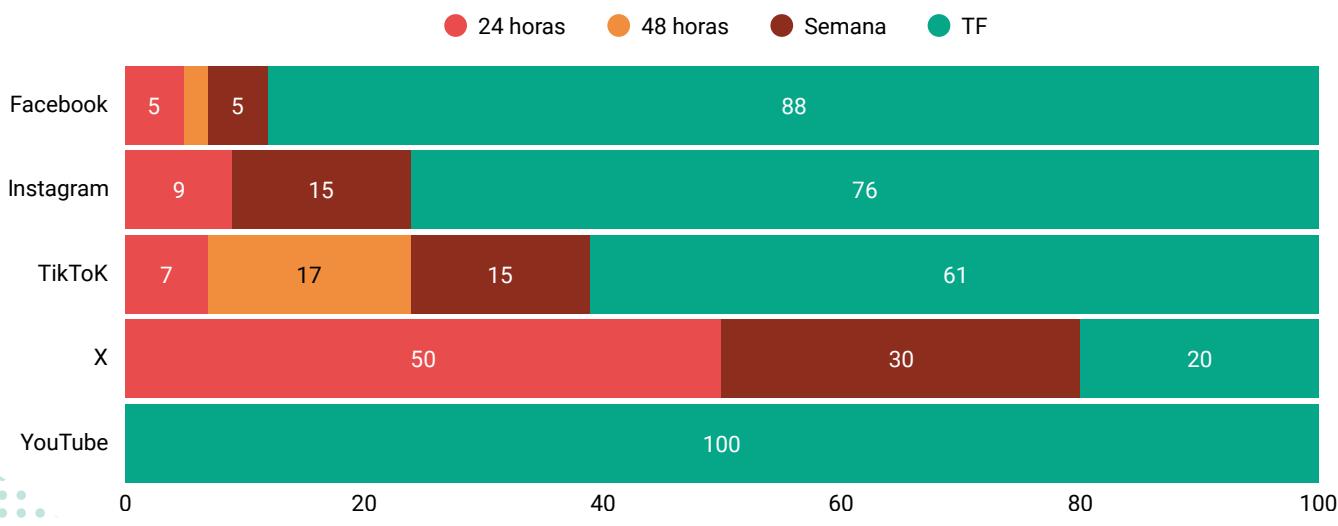
A su vez, mediante la vía *trusted flagger*, empleada este trimestre en 493 ocasiones, las plataformas han retirado 130 publicaciones más, lo que supone un 24% adicional.

A continuación, se muestra el gráfico de contenidos retirados por cada plataforma.



Evolución en la retirada de contenidos

Al analizar los contenidos retirados en base a las notificaciones realizadas a cada plataforma, X, a pesar de su baja tasa de retirada, ha eliminado el 50% de sus casos en un plazo inferior a las 24 horas. En el caso de YouTube, en esta ocasión, solo ha sido efectiva cuando se ha empleado la vía *trusted flagger*.



[Consulta aquí los datos de monitorización en tiempo real](#)

El discurso de odio en el fútbol

Durante el primer trimestre de 2025, el Sistema Faro detectó que un 10% del total del discurso de odio en redes sociales se originó en el ámbito deportivo. Estos discursos han sido dirigidos principalmente a futbolistas a través de insultos y expresiones de deshumanización basadas en estereotipos, aunque también hacia las personas musulmanas por la práctica del Ramadán.

Por otro lado, la iniciativa de LALIGA para medir y combatir los discursos de odio en redes sociales, a través del monitor MOOD que sirve para:

- Monitorizar el discurso de odio en sus diversas tipologías y observar su evolución a lo largo de la temporada de los partidos de LALIGA.
- Impactar generando contenido relevante para las aficiones y los clubes.
- Educar y concienciar ayudando a las organizaciones deportivas a sensibilizarse y tomar medidas contra la violencia en el deporte.
- Promover valores impulsando la tolerancia, el respeto y la inclusión como pilares dentro de las competiciones.

Durante el primer trimestre del año, la jornada con peor nota es la jornada 24, situándose en un 7,4. Por el lado positivo, destaca la nota de la jornada 29, con un 5,1. A continuación, se muestran los datos del monitor MOOD.

5,1

Mood de la jornada

530,2K

Menciones

163 mil

Usuarios

Mood de la jornada: es una puntuación del 1 al 10 que mide el grado de violencia discursiva en redes sociales alrededor de la competición bajo el amparo de La Liga Nacional de Fútbol Profesional.

Estos datos reflejan la persistencia de expresiones discriminatorias en espacios públicos de alta visibilidad hacia las personas de origen extranjero.