



Boletín mensual de monitorización de discurso de odio

Nota de prensa

La monitorización del OBERAXE logra con las plataformas un nuevo hito al superar el 62% de mensajes racistas eliminados en diciembre

- El Observatorio contra el Racismo y la Xenofobia (OBERAXE) reportó 28.978 mensajes con discurso de odio a las plataformas, la cifra más baja del año, que demuestra la eficacia del grupo de trabajo contra el odio
- Se incrementa en 11 puntos porcentuales la tasa de retirada respecto al boletín de noviembre, lo que constata un refuerzo significativo de los mecanismos de moderación y respuesta
- Todas las plataformas aumentan su porcentaje de retirada. TikTok, con un 86%, sigue siendo la más eficaz, seguida de X, que ha retirado un 67% de los mensajes denunciados
- “Este es el resultado de un intenso trabajo, día a día y mano a mano, con las diferentes redes sociales”, ha asegurado el director del OBERAXE, Tomás Fernández Villazala

Madrid, 23 de enero de 2026.- El mes de diciembre ha marcado un nuevo récord en la eliminación de mensajes de contenido racista y xenófobo en redes sociales. Durante el último mes del año el Observatorio Español del Racismo y la Xenofobia (OBERAXE), dependiente de la Secretaría de Estado de Migraciones del Ministerio de Inclusión, Seguridad Social y Migraciones, registró la mayor tasa de retirada de comentarios al conseguir que las plataformas suprimieran un 62% de los mensajes notificados -frente al 51% que se suprimieron en noviembre-. De ellos, un 14% se retiraron en las primeras 24 horas, mientras que un 38% fueron eliminados tras notificarse como institución o verificador confiable (*trusted flagger*).



Nota de prensa

Así se desprende de los datos de monitorización del discurso de odio recogidos en el recientemente publicado boletín de diciembre y realizado a través del sistema FARO (Filtrado y Análisis de Odio en Redes Sociales). A lo largo del mes de diciembre se monitorizaron cinco redes sociales: TikTok, X, Facebook, Instagram y Youtube. En ellas se detectaron 28.978 mensajes, cifra que eleva a 808.176 el total de mensajes detectados a lo largo del año 2025. El dato del mes de diciembre supone además un importante descenso con relación a los mensajes detectados en noviembre, con 10.076 comentarios menos detectados respecto al mes anterior.

Como ha afirmado el director del OBERAXE, Tomás Fernández Villazala, “este es el resultado de un intenso trabajo en el día a día, mano a mano con las distintas redes sociales”, al tiempo que ha destacado que “hemos conseguido un 62% de retirada de los mensajes, lo que nos ayuda cada día a lograr una sociedad más justa, respetuosa y multicultural”. “Y seguimos mejorando en los datos cada mes, lo que nos hace ver que estamos en el buen camino”, ha añadido Fernández Villazala.

TikTok la plataforma más eficaz en la retirada

La plataforma TikTok sigue liderando la lucha contra los mensajes de odio, ya que sigue aumentando su tasa de retirada hasta el 86% de los contenidos reportados, frente al 79% que eliminó en noviembre. Le sigue la plataforma X, con un porcentaje de eliminación del 67%, Facebook con un 42% de comentarios eliminados, Instagram, el 38% y Youtube, con un 32% de retirada de contenidos denunciados.

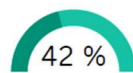
Las redes sociales retiraron el 24% de los contenidos notificados como usuario normal -no como institución verificable-, un dato que supone un aumento notable respecto a noviembre, cuando la retirada por esta vía se situó en el 9%. En relación con el plazo de efectividad, el 14% ha sido retirado en las primeras 24 horas, el 4% en 48 horas y el 6% restante a lo largo de la primera semana tras la notificación. Esto pone de relieve diferencias en la efectividad de las distintas vías de verificación, y demuestra la necesidad de reforzar los mecanismos de moderación y mejorar la eficacia de los reportes realizados como usuarios normales, no institucionales.



Nota de prensa



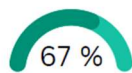
TikTok retiró el 86% de los contenidos notificados; de ellos, el 48% fueron retirados transcurridas las 24 horas desde su notificación.



Facebook ha retirado el 42% de los contenidos notificados, todos ellos a través de la vía *trusted flagger*.



Instagram tuvo una tasa de retirada del 38%, retirando el 50% de los contenidos en un plazo de 48 horas.



X ha retirado el 67% de los contenidos notificados, de los cuales el 20% fue retirado en el plazo de una semana desde su notificación.



YouTube presenta una tasa de retirada del 32%, lo que indica un incremento de 4 puntos porcentuales respecto a noviembre (28%).

Un 70% de contenidos vinculados con personas del norte de África

La hostilidad en el mes de diciembre se concentró una vez más en el grupo diana de personas procedentes del norte de África. Un 70% de los contenidos analizados se referían a estas personas, lo que supone un descenso de seis puntos porcentuales respecto al mes anterior. Además, los comentarios referidos hacia personas musulmanas apenas suponen una variación respecto al mes anterior con un 18%. Mientras que los afrodescendientes protagonizan un 9% de los mensajes vinculados con discurso de odio.

Atendiendo al tipo de contenido, los comentarios que deshumanizan a las personas extranjeras siguen siendo los más frecuentes, a pesar de que han experimentado un descenso significativo pasando de un 68% en noviembre a un 46% en diciembre, pero siguen siendo los más frecuentes y contribuyen a la normalización de las actitudes hostiles.

Además, se duplica el porcentaje de mensajes que presentan al grupo de las personas extranjeros como amenaza y fomentan la percepción de estas personas como un peligro ya que alcanzan un 28% frente al 13% del mes anterior. Los mensajes que incitan a la expulsión suponen un 14%, los que incitan a la violencia,



un 6% y los que alaban a quienes fomentan discursos hostiles suponen un 5%, todos ellos suponen un repunte de 3 puntos respecto al mes anterior lo que implica un incremento de mensajes hostiles y que legitiman la violencia contra estas personas. En el 92% de los mensajes el contenido es además agresivo explícito, lo que refleja altos niveles de hostilidad y contribuye a la polarización social.

Bondi Beach y el desalojo de Badalona

Los episodios que más mensajes eliminados por odio hubo en redes sociales estaban vinculados principalmente con la inseguridad ciudadana, con un 57% del total. Esto evidencia la instrumentalización de la inseguridad ciudadana para reproducir narrativas xenófobas y racistas que consolidan estereotipos basados en datos falsos que acaban impactando en la opinión pública.

Entre los episodios más frecuentes destacan el conflicto armado (20%) en relación con la guerra entre Israel y Palestina, que fue detonante de comentarios con contenido antisemita e islamófobo con mensajes que justifican la violencia. También destacan en relación con el terrorismo (16%) los mensajes de odio retirados tras los atentados en Bondi Beach, en Australia.

Respecto al ámbito económico, que concentra un 14% de los mensajes detectados, cabe destacar el odio generado tras el desalojo en Badalona de una antigua escuela de secundaria en la que residían 400 inmigrantes.

Precisamente para seguir mejorando en la eliminación del discurso en redes sociales, se puso en marcha el grupo de trabajo para la eliminación del discurso de odio en redes sociales que se reunió el pasado 14 de enero y en el que tanto las plataformas como la Secretaría de Estado de Migraciones mostraron su predisposición para seguir avanzando en este camino de la mano.

Más de cinco años monitorizando las principales redes sociales

La monitorización del discurso de odio en las redes sociales que realiza el OBERAXE desde 2020 consiste en la búsqueda, recopilación, análisis y notificación a las plataformas, de contenidos de discurso de odio con motivación racista, xenófoba, islamófoba, antisemita y antigitana 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.



Nota de prensa

En el año 2024, 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 a través del nuevo 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.