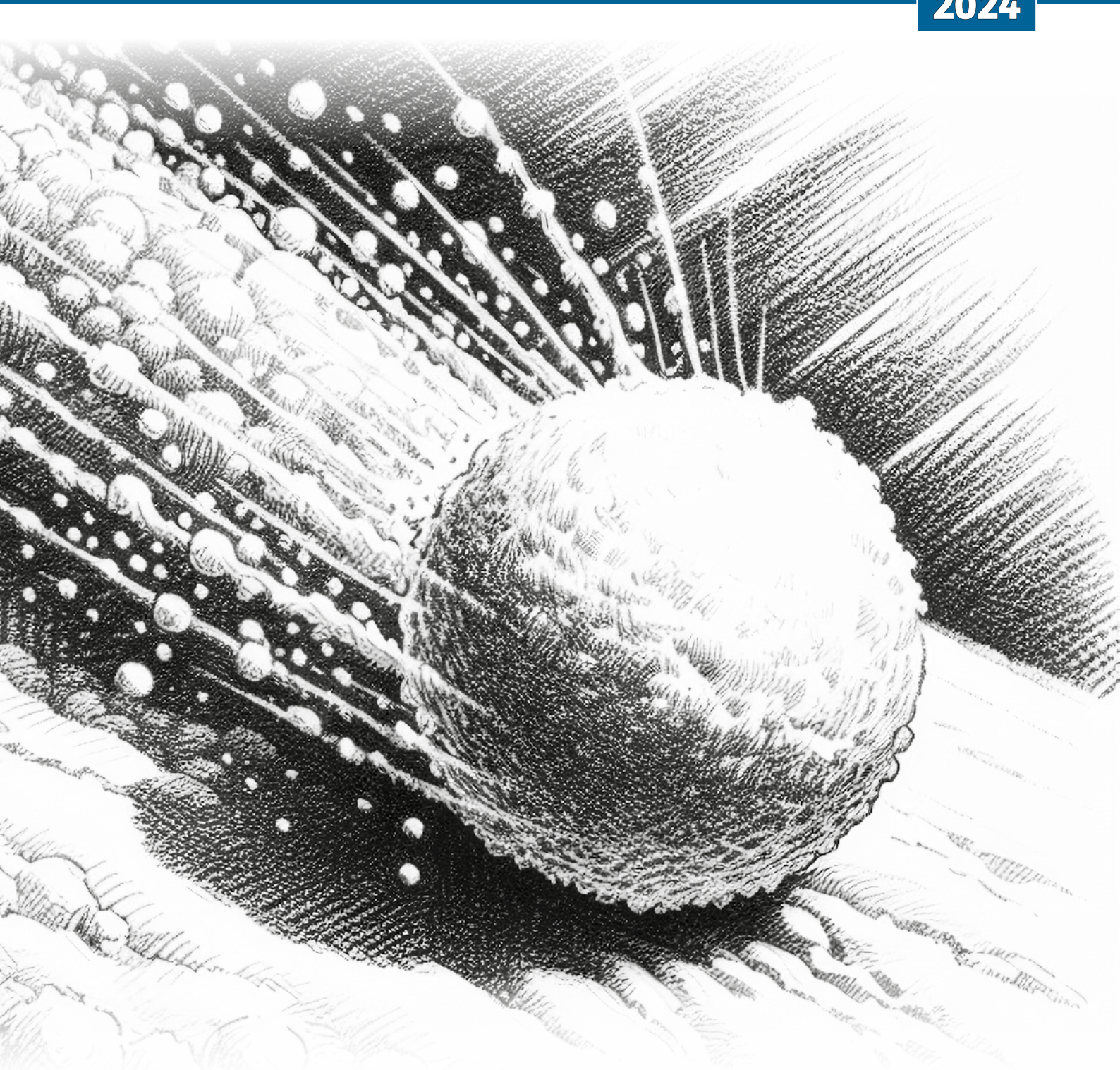


INFORME ANUAL

MONITORIZACIÓN DEL DISCURSO DE ODIO EN REDES SOCIALES

Contenidos considerados de odio racista y/o xenóforo, islamóforo, antisemita y antigitano

2024



GOBIERNO DE ESPAÑA

MINISTERIO DE INCLUSIÓN, SEGURIDAD SOCIAL Y MIGRACIONES

SECRETARÍA DE ESTADO DE MIGRACIONES



Cofinanciado por la Unión Europea



Catálogo de publicaciones de la Administración General del Estado

<https://cpage.mpr.gob.es>

© Ministerio de Inclusión, Seguridad Social y Migraciones.

Madrid, 2025

Autor/es: Observatorio Español del Racismo y la Xenofobia (OBERAXE)

Edita: Observatorio Español del Racismo y la Xenofobia

Distribuye: Observatorio Español del Racismo y la Xenofobia

Calle Agustín de Betancourt, 11, séptima planta. 28003 Madrid

Correo electrónico: oberaxe@inclusion.gob.es

Web: <https://www.inclusion.gob.es/oberaxe/es/index.htm>

NIPO: 121-24-009-4

Diseño: DISEÑO GRÁFICO GALLEGO Y ASOCIADOS, S. L.

Maquetación: CYAN, Proyectos Editoriales, S.A.

Cómo citar: Observatorio Español del Racismo y la Xenofobia [OBERAXE] (2025). *Informe Anual de Monitorización del Discurso de Odio en Redes Sociales 2024*. Recuperado de: https://www.inclusion.gob.es/oberaxe/es/publicaciones/documentos/documento_0164.htm

1. ÍNDICE

1	1. INTRODUCCIÓN.....	5
2	2. OBJETIVO.....	8
3	3. METODOLOGÍA.....	10
4	4.RESULTADOS.....	12
5	5.CONCLUSIONES.....	29
6	6.ANEXO I: Ejemplos de discurso de odio.....	31

Índice de tablas

Tabla 1. Porcentaje de contenidos retirados según el tiempo transcurrido desde la notificación y según plataforma de internet, 2024.....	13
Tabla 2. Distribución de la tipología del discurso de odio notificado.....	18
Tabla 3. Distribución de la tipología del discurso de odio (%) en cada grupo diana.....	19

Índice de gráficos

Gráfico 1. Porcentaje de comunicaciones a cada plataforma de internet.....	12
Gráfico 2. Porcentaje de contenidos retirados según el tiempo transcurrido desde la notificación al conjunto de las plataformas monitorizadas, 2024.....	13
Gráfico 3. Porcentaje de discurso de odio dirigido a cada grupo diana.....	15
Gráfico 4. Evolución de los principales grupos diana en 2024.....	16
Gráfico 5. Prevalencia de los grupos diana de discurso de odio en cada red social.....	17
Gráfico 6. Distribución de la tipología del discurso de odio (%) en cada grupo diana.....	18
Gráfico 7. Distribución de la tipología del discurso de odio según red social.....	19
Gráfico 8. Tipología del discurso de odio según reacción de las plataformas respecto a la retirada.....	20
Gráfico 9. Frecuencia de la expresión del discurso de odio.....	21
Gráfico 10. Distribución de la expresión del discurso de odio según el grupo diana al que se dirige.....	21
Gráfico 11. Distribución del tipo de expresión del discurso de odio según la plataforma de internet.....	22
Gráfico 12. Distribución del discurso de odio notificado a las plataformas según su vinculación con un episodio prototípico.....	24
Gráfico 13. Evolución de los principales episodios prototípicos en 2024.....	25
Gráfico 14. Distribución de episodios prototípicos relacionados con el discurso de odio según grupo diana.....	26
Gráfico 15. Distribución de la expresión del discurso de odio según el episodio prototípico.....	27



1. INTRODUCCIÓN

El auge de las redes sociales ha transformado profundamente la manera en que las personas se comunican, acceden a la información y participan en la esfera pública. Esta transformación también ha traído consigo nuevos retos en materia de convivencia, respeto y protección de los derechos fundamentales. En particular, el discurso de odio en entornos digitales representa un fenómeno creciente que impacta negativamente sobre la cohesión social, la seguridad de los colectivos vulnerables y de la sociedad en general y la calidad del debate público.

Ante esta realidad, la necesidad de contar con herramientas institucionales que permitan observar y analizar, y que ofrezcan información para combatir el discurso de odio en redes sociales, se ha hecho cada vez más evidente.

En respuesta a este desafío, 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, implementó en el año 2020 un sistema de monitorización del discurso de odio en redes sociales que permite analizar las características de este tipo de discurso, y sus resultados

se presentan periódicamente en los boletines publicados en la [web del OBERAXE](#), inicialmente de forma bimestral y, desde septiembre de 2024, con una frecuencia mensual.

Asimismo, en 2023 se publicó el primer informe anual sobre los resultados de la monitorización del discurso de odio, que constituyó un paso fundamental en la consolidación de un enfoque sistemático y basado en evidencias para abordar este fenómeno. En dicho informe se presentaron las bases conceptuales, metodológicas e institucionales empleadas para llevar a cabo un seguimiento riguroso y sostenido del discurso de odio, proporcionando por primera vez un análisis anual detallado de su alcance, temáticas principales y patrones recurrentes.

El presente informe, correspondiente al año 2024, da continuidad a esa línea de trabajo, manteniendo el enfoque general del año anterior, a la vez que introduce las mejoras metodológicas y ajustes técnicos que permiten afinar el análisis y mejorar la detección de fenómenos emergentes.

Uno de los hitos más destacados para el OBERAXE en 2024 ha sido la firma, el 24 de octubre, del convenio entre el Ministerio de Inclusión, Seguridad Social y Mi-

graciones y LALIGA, que ha permitido el desarrollo del Sistema FARO.

El Sistema FARO es la nueva metodología que va a emplear el OBERAXE, a partir de marzo de 2025, para la identificación y análisis en tiempo real de los contenidos de discurso de odio con motivación racista, xenófoba, islamófoba, antisemita y antigitana, y que permite incorporar la inteligencia artificial avanzada para la monitorización de las redes sociales, aumentando notablemente el número de contenidos potencialmente constitutivos de discurso de odio identificados.

El Sistema FARO incorpora el uso de la tecnología de la inteligencia artificial, entrenada en el Monitor para la Observación del Odio en el Deporte (MOOD) de LALIGA, al conocimiento y la experiencia que atesora el OBERAXE en la monitorización del discurso de odio en las redes sociales.

El Sistema FARO es la conjunción del uso de dos herramientas (Monitor FARO y ALERTODIO) aunadas a una nueva metodología de trabajo que atañe tanto a la identificación de contenidos como al análisis y presentación de los resultados a través de un monitor de visualización de datos en tiempo real. El monitor de visualización de datos del Sistema FARO está disponible para consulta de toda la ciudadanía en el nuevo portal web del OBERAXE.

La monitorización sistemática del discurso de odio en redes sociales que realiza el OBERAXE, se hace en colaboración estrecha con la Fiscalía General del Estado y las Fuerzas y Cuerpos de Seguridad del Estado. Esta cooperación se canaliza a través de los Equipos de Respuesta a los Delitos de la Guardia Civil (REDO) y los equipos especializados en Extremismo Violento y Odio (EVO) de la Policía Nacional, enfocándose en la detección e investigación de incidentes y delitos de odio complejos.

Además, la actividad del OBERAXE se nutre de proyectos cofinanciados por la Comisión Europea como el [proyecto CISDO](#), realizado en colaboración con la Oficina Nacional de Lucha Contra los Delitos de Odio (ONDOD) del Ministerio del Interior. Este proyecto busca mejorar las capacidades policiales a nivel nacional y local para prevenir, identificar y responder a los incidentes xenófobos y racistas, así como a los delitos de odio, y para proporcionar una mejor asistencia a las víctimas; el [proyecto SCORE](#) cuyo objetivo ha sido crear una coalición de ciudades y entidades locales europeas para la promoción del deporte inclusivo, así como la prevención y la lucha contra el racismo, la xenofobia y las formas conexas de intolerancia en el ámbito deportivo; y el proyecto [REAL UP](#), destinado a mejorar las capacidades de las autoridades estatales para identificar, analizar, supervisar y evaluar el discurso de odio en línea con el

fin de desarrollar y fortalecer las estrategias de contranarrativa (discurso upstander) contra el discurso de odio motivado por el racismo, la xenofobia, la islamofobia, el antisemitismo y el antigitanismo.

Por otro lado, la reciente entrada en vigor plena del [Reglamento \(UE\) 2022/2065, conocido como Ley de Servicios Digitales \(DSA\)](#), a partir del 17 de febrero de 2024, representa un paso significativo para regular la responsabilidad de los prestadores de servicios digitales, incluyendo las plataformas digitales, motores de búsqueda y otros proveedores online. Este reglamento establece obligaciones claras para que estas plataformas garanticen un espacio digital seguro, transparente y competitivo, protegiendo los derechos fundamentales de los usuarios y fomentando la innovación en el mercado único europeo.

En el plano europeo, también, el 20 de enero de 2025 se lanzó la versión renovada del [Código de Conducta para la lucha contra la incitación ilegal al odio en internet+](#), firmado con la Comisión Europea y por las principales plataformas digitales como Meta, X, YouTube, TikTok, LinkedIn y Microsoft, entre otras. Este Código refuerza los compromisos voluntarios asumidos en 2016 y complementa el marco legal establecido por la ley de servicios digitales (DSA), promoviendo políticas claras que prohíben el discurso de odio, mecanismos efectivos de notificación para usuarios europeos, revisión diligente de reportes y transparencia en los procesos de moderación humana y automatizada. Asimismo, se establece un sistema de cooperación con organizaciones de la sociedad civil especializadas, que permite compartir conocimiento y mejorar la identificación y análisis del discurso de odio en línea.

En España, la Comisión Nacional de los Mercados y la Competencia (CNMC) fue designada coordinadora nacional para la aplicación de la Ley de Servicios Digitales el 24 de enero de 2024, asumiendo un papel clave como único punto de contacto y responsable de supervisar la correcta implementación de la DSA a nivel nacional.

Además, la CNMC mantiene su compromiso, participando activamente en los grupos de trabajo de la Junta Europea de Servicios Digitales y promoviendo el diálogo con instituciones nacionales, incluida su colaboración con el OBERAXE y su participación en el "Acuerdo para Cooperar Institucionalmente contra el racismo, la xenofobia, la LGTBIfobia y otras formas de intolerancia". El OBERAXE aspira a ser nominado "*trusted flagger*" por la CNMC a lo largo del año 2025, dado su trabajo y experiencia en la detección y lucha contra los discursos de odio y su implicación en el ámbito internacional y europeo, entre otros, en el Grupo de Alto Nivel de Delitos y Discurso de Odio de la Comisión Europea.

El OBERAXE valora muy positivamente el reciente nombramiento de la Fiscal María Teresa Verdugo Moreno como Autoridad Independiente para la Igualdad de Trato y la no Discriminación, dando cumplimiento a la Ley 15/2022, integral para la igualdad de trato y la no discriminación, ya que se facilitará la lucha contra muchos de esos contenidos discriminatorios que no alcanzan la consideración para ser tramitados por la vía penal. También se podrá avanzar en la tramitación más ágil de sanciones administrativas para los perpetradores, que desalentarán este tipo de conductas. Del mismo modo que las actuaciones del Consejo Superior de Deportes, en virtud de la Ley 19/2007 contra la Violencia, el Racismo, la Xenofobia y la Intolerancia en el Deporte, que ha tenido algunas sentencias ejemplarizantes.

España ha demostrado ser referente de la lucha contra el discurso de odio por el trabajo de distintas instituciones, entre ellas OBERAXE. De hecho, la monitorización continua y colaboración con actores públicos y privados ha sido reconocida, entre otros, por el Secretario General adjunto del Consejo de Europa, Bjørn Berge, que destacó en diciembre de 2024 la relevancia y el impacto del trabajo desarrollado por el OBERAXE en este ámbito.

Además, en octubre de 2024 se llevó a cabo unas jornadas sobre contranarrativas y narrativas alternativas organizadas por el OBERAXE en colaboración con el Consejo de Europa, a las que asistieron una veintena de organizaciones de la sociedad civil y otros actores institucionales que trabajan a favor de la inclusión de personas inmigrantes y en contra de la discriminación.

El objetivo principal de las jornadas fue dotar a las organizaciones participantes de conocimientos y herramientas para la formulación de contranarrativas y narrativas alternativas que permitan combatir los discursos de odio racistas y xenófobos y sensibilizar a través de campañas de comunicación, con un enfoque basado en los derechos humanos. En ellas se abordaron conceptos como el impacto del discurso de odio y el análisis de las narrativas discriminatorias en el contexto de España.

En conclusión, la situación del discurso de odio en redes sociales es compleja, requiere esfuerzo y recursos, así como abordajes desde distintos ámbitos para combatirla: legislación penal y administrativa, estrategias de contranarrativa, sensibilización y formación, monitorización y análisis de la situación.

La integración de recursos tecnológicos como la inteligencia artificial, el fortalecimiento de las instituciones nacionales y europeas, y la cooperación con plataformas y la sociedad civil, conforman los pilares sobre los que se debe sustentar la respuesta efectiva ante este desafío. La monitorización del discurso de odio no solo permite diagnosticar la situación, sino también contribuir con información para el diseño de políticas y estrategias que contribuyan a garantizar un entorno digital libre de discriminación, intolerancia y hostilidad hacia las personas de origen extranjero, favoreciendo así su inclusión en la sociedad y la cohesión social.



2. OBJETIVO

El objetivo general del sistema de monitorización del OBERAXE es identificar, caracterizar y evaluar la presencia del discurso de odio en redes sociales, con el propósito de:

- Contribuir al conocimiento empírico sobre su magnitud y evolución.
- Detectar tendencias emergentes y focos temáticos de hostilidad.
- Informar y orientar las políticas públicas de prevención, sensibilización e intervención.
- Fortalecer la colaboración institucional y con actores del ámbito digital, académico y de la sociedad civil.

El sistema abarca tanto expresiones de odio explícitas como formas más sutiles, implícitas o codificadas, reconociendo que los discursos discriminatorios adoptan múltiples formatos, grados de intensidad y niveles de visibilidad.

La monitorización que realiza OBERAXE se centra en la búsqueda, recopilación, análisis y notificación a las plataformas digitales de contenidos que constituyen

discurso de odio con motivaciones racistas, xenófobas, islamófobas, antisemitas y antigitanas. Estos contenidos pueden ser constitutivos de delito, infringir normativas administrativas o vulnerar las normas de uso de las propias plataformas.

El ámbito de actuación abarca exclusivamente aquellos discursos dirigidos a personas o grupos por razón de su origen étnico, nacional o religión. Debido a sus particularidades, se presta especial atención a grupos vulnerables como personas inmigrantes, niños, niñas y jóvenes no acompañados, y personas refugiadas.

Objetivos específicos

Los objetivos específicos de la monitorización son dos:

- 1. Evaluar la respuesta de las plataformas:** Se analiza cómo las plataformas digitales gestionan la retirada de contenidos de discurso de odio ilegal que se les notifican, en cumplimiento de los compromisos ad-

quiridos en el marco del Código de Conducta de la Unión Europea y las obligaciones establecidas por la Ley de Servicios Digitales (DSA).

La moderación en las plataformas se basa en dos pilares fundamentales:

- La retirada de contenidos ilegales según la legislación nacional de los Estados miembros de la UE, conforme a lo dispuesto en la DSA.
- La eliminación de contenidos que infrinjan las normas de uso internas de cada plataforma, una acción voluntaria que responde a sus propios compromisos, incluida la adhesión al Código de Conducta europeo.

2. Analizar el discurso de odio y detectar tendencias:

Además de la notificación y evaluación de respuestas, se realiza un análisis detallado del discurso de odio racista, xenófobo, islamófobo, antisemita y antigitano en España, contribuyendo así a una mejor comprensión de la situación y facilitando la orientación del diseño y la implementación de políticas públicas.

Cabe resaltar que en el OBERAXE se mantiene una colaboración activa con las plataformas digitales, estableciendo canales de comunicación continuos que incluyen reuniones periódicas para el intercambio de información, análisis y discusión sobre la identificación y moderación de los contenidos de odio.

3. METODOLOGÍA



Este informe presenta los resultados del ejercicio de monitorización del discurso de odio en redes sociales desarrollado por el OBERAXE, correspondientes al año 2024. Da continuidad al trabajo iniciado en años anteriores, especialmente al desarrollado en el Informe de monitorización de 2023, en el que se presentaron los fundamentos conceptuales, los objetivos y la estructura metodológica del sistema.

El sistema de monitorización desarrollado por OBERAXE se basa en una metodología mixta, que permite abordar la complejidad del fenómeno desde una perspectiva multidimensional, incorporando elementos cuantitativos, lingüísticos, contextuales y socioculturales.

La metodología aplicada se basa íntegramente en una búsqueda manual y sistemática de contenidos, realizada de forma diaria por un equipo de monitorizadores. Este equipo lleva a cabo la observación directa de perfiles abiertos en redes sociales, así como el seguimiento de debates, noticias de actualidad, publicaciones virales y el

uso de un glosario propio de términos y expresiones actualizado periódicamente. Esta estrategia permite captar no solo los mensajes abiertamente hostiles, sino también formas más sutiles o codificadas de discurso de odio.

La complejidad de los discursos discriminatorios y su dependencia del contexto sociopolítico hacen que la revisión manual resulte fundamental para garantizar un análisis riguroso, capaz de identificar matices, referencias culturales, eufemismos o elementos interseccionales que podrían pasar desapercibidos con metodologías automatizadas.

Todos los contenidos identificados son registrados de forma sistemática en la aplicación ALERTODIO, desarrollada en colaboración con la Universidad Politécnica de Valencia. Esta herramienta facilita un registro homogéneo, detallado y estructurado de la información, permitiendo la anotación de variables clave ¹ como la motivación del discurso, el contexto de publicación, el tipo de contenido, el lenguaje que se emplea y el grupo diana al que se dirige.

1. Definición de variables: ver capítulo 3.2.1. del Informe Anual de Monitorización del discurso de odio en Redes Sociales 2023

Una vez registrados los contenidos que pudieran constituir discurso de odio ilegal o que infrinjan las normas comunitarias de las plataformas digitales, se procede a su notificación a las mismas mediante un procedimiento escalonado. A continuación, a partir de la notificación inicial realizada como usuario normal, se procede a un seguimiento sistemático de la respuesta de las plataformas (si retiran o no el contenido denunciado), con revisiones a las 24 horas, 48 horas y al cabo de una semana. En caso de que la plataforma no haya retirado el contenido tras este período, se vuelve a notificar el

contenido, esta vez como *trusted flagger*² o comunicante fiable. Este proceso permite valorar la respuesta de las plataformas y evaluar su compromiso con la moderación de contenidos discriminatorios.

Para un desarrollo más detallado del sistema de identificación, clasificación y notificación de contenidos, así como de los criterios empleados, se remite al lector al [Informe Anual de Monitorización del discurso de odio en Redes Sociales 2023](#), donde se describe en profundidad la metodología del OBERAXE para la monitorización del discurso de odio, y que se mantiene en este informe.

2. Definición de *trusted flagger*: son individuos o entidades acreditadas especialmente por el prestador de servicios de alojamiento de datos por entender que cuentan con la cualificación o conocimiento necesario para notificar contenido de discurso de odio, al dedicar su actividad a cuestiones relacionadas con la lucha contra la discriminación y atesorar la experiencia necesaria en este ámbito



4. RESULTADOS

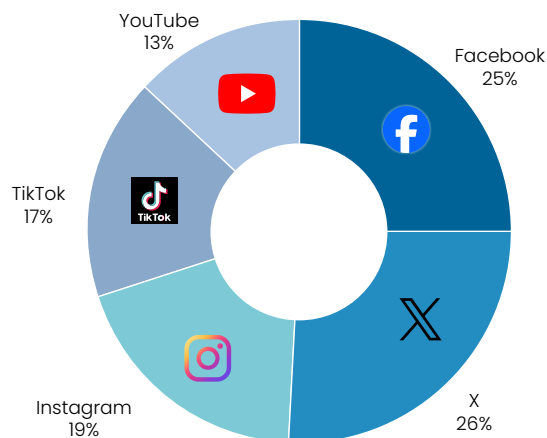
4.1. Contenidos monitorizados y notificados y reacción de las plataformas de redes sociales

En 2024, se han notificado 2.870 contenidos considerados de odio racista, xenófobo, antisemita, antigitano o islamóforo, que podían ser constitutivos de delito, de infracción administrativa o que violan las normas de conducta de las plataformas digitales, a las cinco redes sociales monitorizadas (Facebook, X, Instagram, TikTok y YouTube).

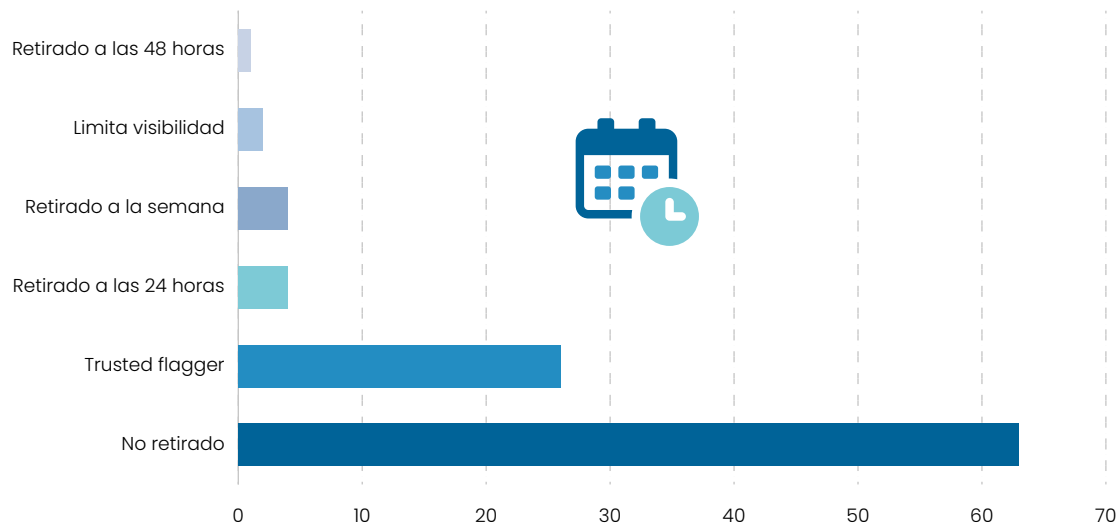
En la distribución de comunicaciones realizadas a cada plataforma (gráfico 1) hay un predominio de X con 758 casos (26% del total), seguido de Facebook con 727 casos (25%), Instagram con 538 (19%), TikTok con 478 (17%) y YouTube con 369 (13%). Este desigual volumen de notificación de contenidos obedece, principalmente, al distinto grado de dificultad para su identificación en cada red social.

Las plataformas retiraron un total de 1.010 contenidos, un 35% de los que les fueron notificados. Del conjunto de contenidos comunicados, únicamente el 9% (272) fueron eliminados cuando la notificación se realizó a través de un perfil de usuario normal, mientras que el 26% (738) fueron retirados tras ser reportados mediante la figura

Gráfico 1. Porcentaje de comunicaciones a cada plataforma de internet



de trusted flagger. Estos datos evidencian una mayor eficacia en la retirada de contenidos cuando la notificación se realiza a través de canales oficiales o reconocidos como comunicante fiable.

Gráfico 2. Porcentaje de contenidos retirados según el tiempo transcurrido desde la notificación al conjunto de las plataformas monitorizadas, 2024**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%
Total	35%	4%	1%	4%	26%	65%

No obstante, la tasa de retirada del conjunto de las plataformas es muy baja y, en comparación con el año 2023, ha disminuido en 14 puntos porcentuales.

En cuanto a la tasa de retirada en base a las notificaciones realizadas a cada plataforma, la más eficiente es TikTok, que retiró el 69% del total del contenido que se le notificó, seguida por Instagram (49%), Facebook (29%), YouTube (27%) y X (15%).

4.1.1. Características del contenido retirado a las 24 horas, a las 48 horas y a la semana

La reacción del conjunto de las plataformas en relación con el tiempo de retirada del contenido notificado se muestra en el gráfico 2, donde se observa que la mayoría de los reportes son retirados cuando se notifican como *trusted flagger* (26%). La eficiencia y la rapidez en

la retirada de contenidos a las 24h, a las 48h, a la semana o por la vía de *trusted flagger* se muestra en la tabla 1, y los resultados difieren entre las cinco plataformas, aunque todas ellas son más receptivas a la eliminación de contenido cuando se emplea la vía de *trusted flagger*.

Los datos evidencian una respuesta poco efectiva por parte de las plataformas ante notificaciones realizadas desde perfiles de usuario normal, especialmente en las primeras 24 y 48 horas, periodos relevantes para minimizar el impacto del discurso de odio. La baja tasa de retirada inmediata, solo el 4% en 24 horas, revela deficiencias en los sistemas de moderación. Esta escasa reacción inicial por parte de las plataformas facilita que contenidos que deshumanizan, promueven estigmas o incitan a la violencia permanezcan visibles y circulen ampliamente, afectando especialmente a los grupos diana. Estas dinámicas pueden contribuir a la normalización del discurso de odio online, lo que subraya

la necesidad de reforzar los mecanismos de moderación y respuesta de las plataformas.

De los contenidos que han sido comunicados como usuario normal, TikTok es la red social que más contenido retiró en las primeras 24 horas (15%), seguida de X (4%), Instagram (2%), Facebook (1%) y por último YouTube que solo ha retirado un 0,3% de sus contenidos en 24 horas.

El 36% de los contenidos retirados a las 24 horas contienen lenguaje que deshumaniza o degrada o que expresa agresividad. El grupo diana mayoritario de los contenidos retirados a las 24 horas ha sido el de las personas originarias del norte de África (43% de los casos) y predomina la inseguridad ciudadana como episodio prototípico (40%).

El conjunto de plataformas ha retirado el 1% de los contenidos en el plazo de 48 horas desde su notificación. En los contenidos retirados en este plazo predominan aquellos que deshumanizan o degradan gravemente, en un 53%, y en un 50%, los que promueven el descrédito en base a atributos personales. El episodio prototípico principal sigue siendo la inseguridad ciudadana en 32% en las notificaciones retiradas, y el grupo diana también son las personas originarias del norte de África (65%).

La plataforma que más contenido ha retirado a las 48 horas ha sido TikTok, que ha eliminado el 3% de los contenidos que se le notificaron, seguida de Instagram (2%), X (1%), Facebook (1%) y YouTube (0,3%).

Respecto al contenido retirado en el plazo de una semana, el conjunto de las plataformas ha retirado el 4% de las notificaciones, siendo Instagram la plataforma que más contenido ha retirado en este plazo temporal (7%), seguida de Facebook (5%), TikTok (5%), X (3%) y YouTube (1%). Cabe resaltar que el 28% de los casos eliminados a la semana no están vinculados a ningún episodio prototípico, y el 59% de las comunicaciones retiradas en ese plazo contienen un lenguaje agresivo explícito.

4.1.2 Contenido retirado como Trusted Flagger

La vía de *trusted flagger* continúa consolidándose como el mecanismo más efectivo para la retirada de contenido de discurso de odio por parte de las plataformas. De las 2.870 notificaciones, un 26% fueron eliminadas tras ser comunicadas mediante esta vía, en contraste con el 9% de efectividad observada cuando la notificación se realizó desde un perfil de usuario normal. La diferencia en las tasas de retirada según la vía, es especialmente significativa, evidenciando que las plataformas dan prioridad a los comunicantes fiables.

Al desagregar los datos por plataforma, se aprecian diferencias significativas en el nivel de eficacia en la retirada de contenidos. TikTok es la plataforma más eficiente ante el comunicante fiable, con un 47% de contenidos retirados por esta vía. Le siguen Instagram, con

un 37%; YouTube, con un 25%; Facebook, con un 23%; y, en último lugar, X, que tiene una tasa de retirada del 7% a través de esta vía.

4.1.3 Características del contenido no retirado

El porcentaje de contenido no retirado ha sido del 65% (1.860 casos). Este porcentaje incluye un 2% de notificaciones para las que se ha reducido la visibilidad por parte de la red social X, la cual estableció este mecanismo en 2023 como acción positiva para disminuir el efecto del contenido de odio, que, aunque continúa circulando por la red, es menos visible para las personas usuarias.

A pesar de las normas y mecanismos que establecen las plataformas en el marco del Código de Conducta y la normativa establecida por la 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. El análisis cualitativo de los 1.860 contenidos no retirados muestra lo siguiente:

- En un 38% de los casos se promueve el descrédito hacia atributos personales.
- El 36% deshumaniza o degrada gravemente al grupo diana.
- Un 28% incita a la violencia mediante amenazas directas o indirectas.
- Un 17% incita a la expulsión de personas de origen extranjero.
- 633 casos dirigidos hacia las personas del norte de África no han sido retirados.
- Predomina la narrativa de la vinculación de la inseguridad ciudadana con los grupos diana.

4.2. Características del discurso de odio en las redes sociales

4.2.1. Grupos diana

Uno de los ejes fundamentales en la monitorización del discurso de odio en redes sociales del OBERAXE es analizar a quién se dirige el discurso. Esta identificación permite no solo mapear los patrones de hostilidad hacia las personas de origen extranjero, sino también orientar acciones que prevengan dinámicas de discriminación.

En 2024, los datos recopilados por el OBERAXE revelan que, en casi ocho de cada diez contenidos notificados a las plataformas, el mensaje tenía como destinatario un grupo diana específico (80%), frente a un 20% de mensajes dirigidos a personas concretas. Esta tendencia confirma

la naturaleza estructural del discurso de odio, que tiende a reforzar estereotipos y alimentar actitudes hostiles hacia los diferentes grupos diana.

Al analizar en detalle los grupos diana del discurso de odio, destaca la hostilidad predominante hacia las personas originarias del norte de África, que concentran el 35% del total de discursos de odio identificados. Le siguen las personas africanas y afrodescendientes, con un 24%, y la categoría genérica de personas inmigrantes, que representa un 21%. Este último dato resulta especialmente significativo, ya que evidencia una tendencia a la generalización de la discriminación por el mero hecho de tener otro origen o nacionalidad.

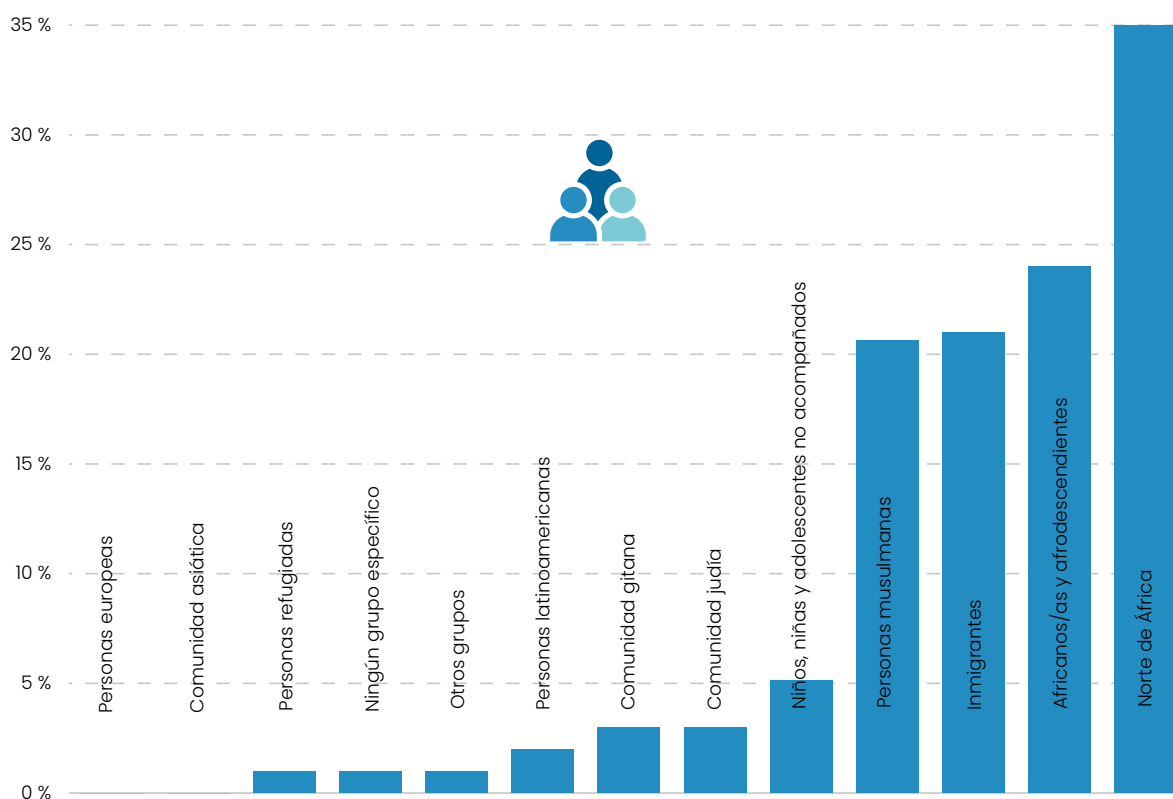
Por otro lado, los discursos islamófobos también mantienen un peso relevante. Las personas musulmanas constituyen el cuarto colectivo más afectado, con un 21% del total. Estas cifras reflejan una continuidad en las dinámicas de estigmatización hacia comunidades con marcadores religiosos visibles, que suelen ser instrumentalizadas mediante narrativas de inseguridad, criminalización y otras formas de exclusión simbólica.

Otro aspecto relevante del análisis es el impacto del contexto geopolítico sobre la evolución del discurso de odio. El conflicto en Oriente Próximo (Israel y Palestina) ha tenido un efecto directo sobre el volumen de mensajes antisemitas en redes sociales. El discurso de odio hacia la comunidad judía se sitúa en el 3%. De forma similar, se ha detectado un aumento de mensajes hostiles hacia las personas musulmanas en relación con este acontecimiento.

También, cabe destacar el discurso de odio dirigido hacia los niños, niñas y adolescentes no acompañados, que registran un 5% de las notificaciones. Este dato refleja cómo los menores no acompañados, especialmente aquellos en situaciones de mayor vulnerabilidad, se convierten en objeto de estigmatización y discriminación, evidenciando la persistencia de prejuicios que los deshumanizan y los exponen a situaciones de riesgo, lo que dificulta su protección y bienestar.

Otros grupos afectados, aunque en menor medida, incluyen la comunidad gitana (3%), personas latinoamericanas (2%), refugiadas (0,52%), asiáticas (0,42%) y europeas (0,31%).

Gráfico 3. Porcentaje de discurso de odio dirigido a cada grupo diana



Si tenemos en cuenta la evolución a lo largo del año de los grupos diana (Gráfico 4) durante los primeros meses de 2024, el discurso de odio mostró fluctuaciones significativas. Destaca especialmente la alta hostilidad hacia la comunidad judía, que alcanzó su punto máximo en enero con un 28%, y tuvo picos secundarios en julio (13%) y octubre (13%). Asimismo, los contenidos hacia personas latinoamericanas y africanas y afrodescendientes fueron persistentes, alcanzando en diciembre valores elevados de 20% y 5%, respectivamente.

Por otro lado, la hostilidad hacia niños, niñas y jóvenes no acompañados presentó un marcado aumento en abril (16%) y en julio (29%), reflejando un foco preocupante sobre este grupo vulnerable en esos meses. Este patrón puede estar vinculado a coberturas mediáticas puntuales o medidas institucionales de gestión migratoria infantil, como por ejemplo la distribución de los menores no acompañados entre las Comunidades Autónomas, que suscitaron reacciones negativas en redes sociales. En ambos meses también se registraron repuntes en el discurso contra otros grupos migrantes, lo que refuerza la hipótesis de una narrativa hostil intensificada en torno a temas migratorios, realizada con agendas políticas.

También se observan altos niveles de rechazo hacia la comunidad gitana con un pico significativo en abril (24%) y junio (17%), así como hacia las personas provenientes del Norte de África, que mostraron picos en agosto (15%) y diciembre (10%).

En relación con la hostilidad dirigida a inmigrantes en general, aunque los porcentajes disminuyeron progresivamente desde enero (13%) a diciembre (4%), este grupo sigue siendo uno de los más afectados de forma constante durante todo el año. Igualmente, la islamofobia y hostilidad hacia las personas musulmanas mostró fluctuaciones, alcanzando su máximo en marzo (13%) y agosto (10%), coincidiendo con la celebración del Ramadán y acontecimientos específicos que se detallan en el apartado de episodios prototípicos.

En referencia al grupo diana al que se dirige el discurso de odio según la plataforma en la que este se difunde de manera predominante, se constata que varía sin que se haya podido identificar ninguna relación causal, aunque puede deberse a las audiencias y dinámicas de cada plataforma (gráfico 5). En la plataforma X los discursos de odio principalmente van dirigidos a las personas del norte de África (30%), seguidos de la comunidad gitana (32%) y las personas africanas y afrodescendientes (29%).

Gráfico 4. Evolución de los principales grupos diana en 2024

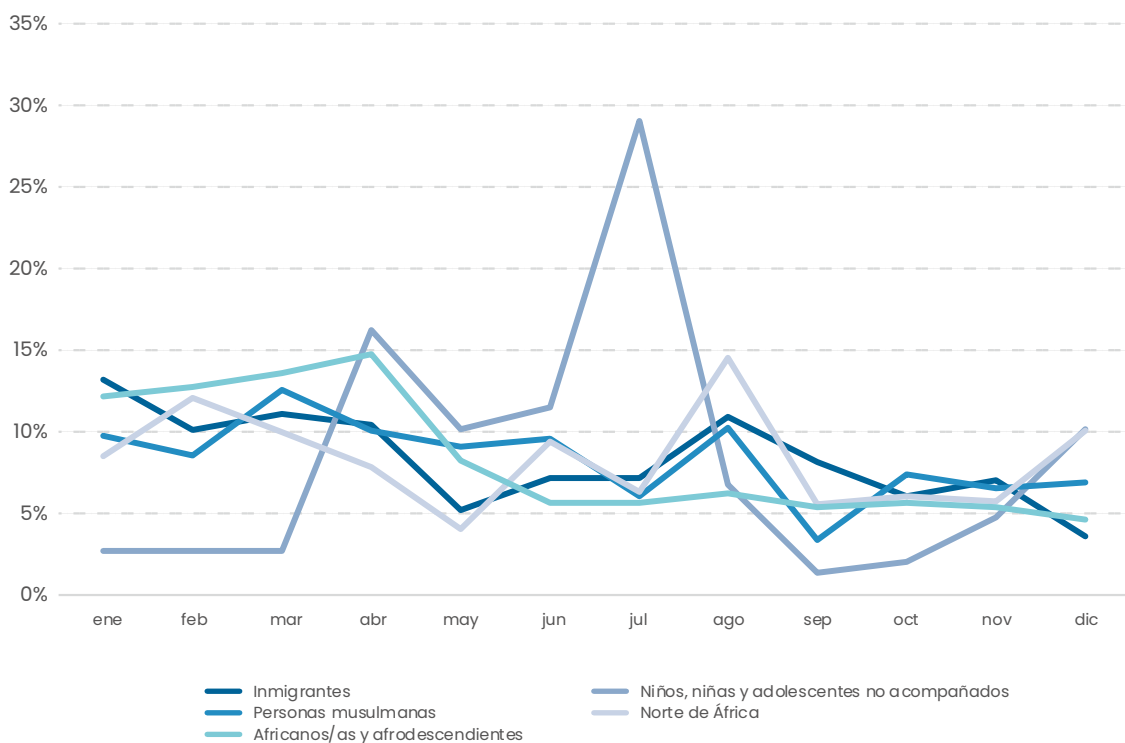


Tabla 2. Distribución de la tipología del discurso de odio notificado

Tipología del discurso	(n)	(%)
Incita a la violencia con amenazas directas o indirectas	843	29
Deshumaniza o degrada gravemente	1058	37
Alaba a quien atenta contra el grupo diana	143	5
Incita a la expulsión del colectivo	430	15
Promueve el descrédito en base a atributos personales	911	32
Presenta al grupo como una amenaza por sus acciones	642	22

la comunidad gitana, prepondera el **descrédito** en el 63% de los contenidos, en base a atributos personales o del colectivo, y que responde a la estigmatización social que sufre este grupo.

También cabe resaltar que en un 32% de los contenidos dirigidos hacia la comunidad judía se **alaba al usuario/a que ejerce el comentario**, dado que la mayoría de estos contenidos hacen alusión al nazismo. En cuanto al colectivo de niños, niñas y adolescentes no acompañados, en un 32% de las notificaciones se presenta al grupo como una **amenaza** para la sociedad.

Por otro lado, al analizar los contenidos de discurso de odio se ha revisado si también se mencionaban otros colectivos vulnerables, además de los grupos de población diana objeto de la monitorización del OBERAXE (mujeres, personas LGTBIQ+, etc.). En la mayoría de las notificaciones esto no ha sucedido, salvo en un 5% donde el discurso va dirigido también hacia las mujeres, y un 0,28% hacia las personas LGTBIQ+.

El Gráfico 7 presenta la tipología de discurso de odio más prevalente en cada plataforma analizada. En YouTube y TikTok predominan los contenidos que incitan

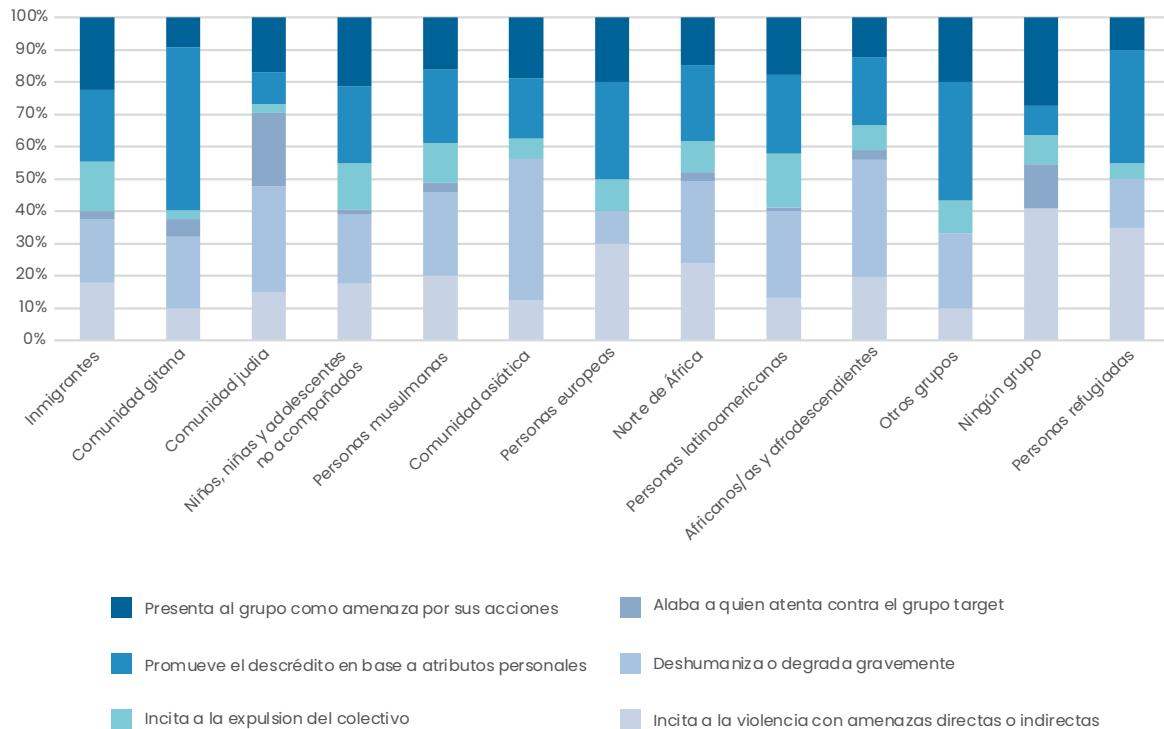
Gráfico 6. Distribución de la tipología del discurso de odio (%) en cada grupo diana

Tabla 3. Distribución de la tipología del discurso de odio (%) en cada grupo diana. Nota: los porcentajes de las filas pueden sumar más de 100 ya que un mismo contenido puede responder a varias tipologías.

	Incita a la violencia con amenazas	Deshumaniza o degrada gravemente	Alaba a quien atenta contra el grupo target	Incita a la expulsión del colectivo	Promueve el descrédito	Presenta al grupo como amenaza
Inmigrantes	26%	29%	4%	22%	32%	32%
Comunidad gitana	13%	28%	7%	3%	63%	11%
Comunidad judía	21%	47%	32%	4%	14%	24%
Niños, niñas y adolescentes no acompañados	27%	32%	3%	22%	36%	32%
Personas musulmanas	30%	38%	4%	18%	34%	24%
Comunidad asiática	17%	58%	0%	8%	25%	25%
Personas europeas	33%	11%	0%	11%	33%	22%
Norte de África	34%	37%	4%	14%	34%	21%
Personas latinoamericanas	19%	38%	2%	23%	34%	25%
Africanos/as y afrodescendientes	26%	48%	4%	10%	28%	16%
Otros grupos	17%	39%	0%	17%	61%	33%
Ningún grupo	53%	0%	18%	12%	12%	35%
Personas refugiadas	47%	20%	0%	7%	47%	13%

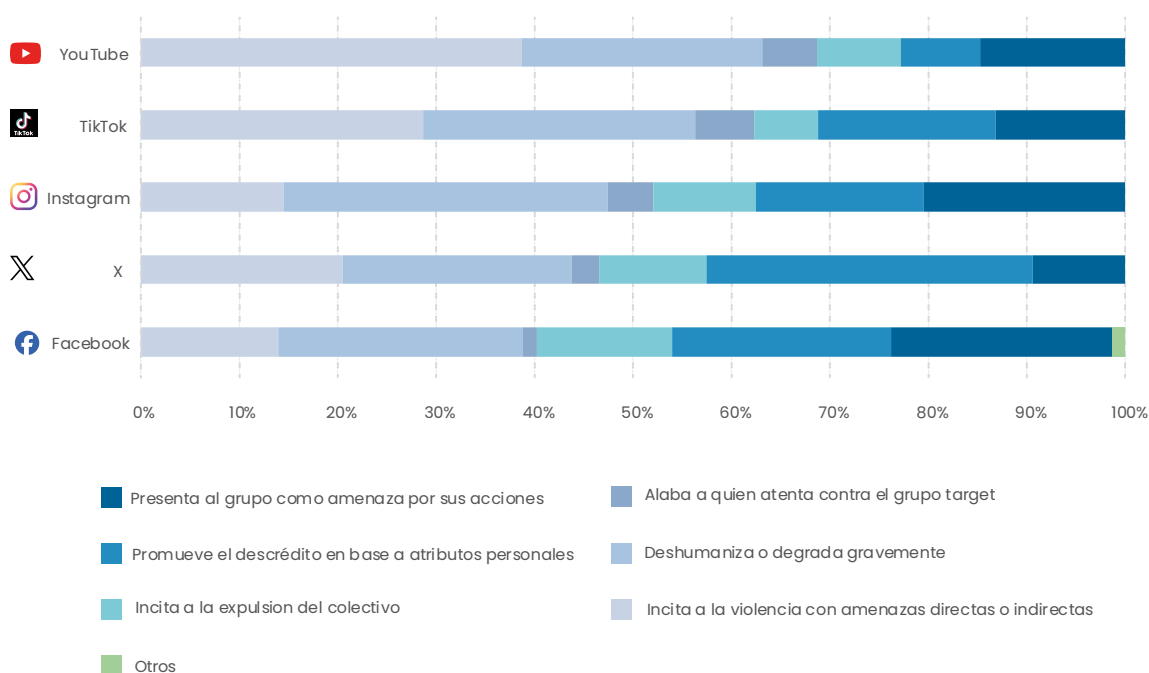
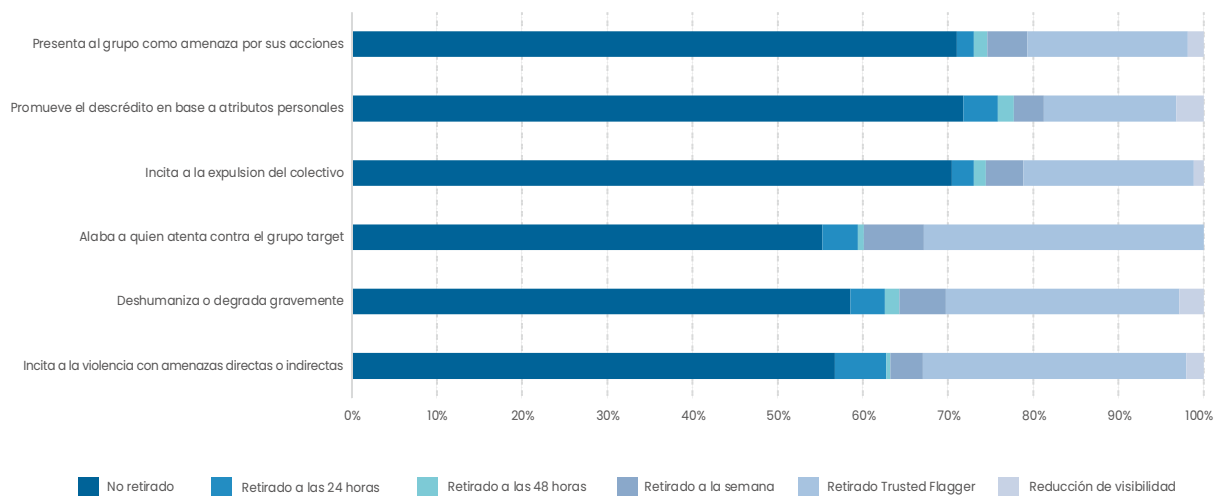
Gráfico 7. Distribución de la tipología del discurso de odio según red social

Gráfico 8. Tipología del discurso de odio según reacción de las plataformas respecto a la retirada

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%.

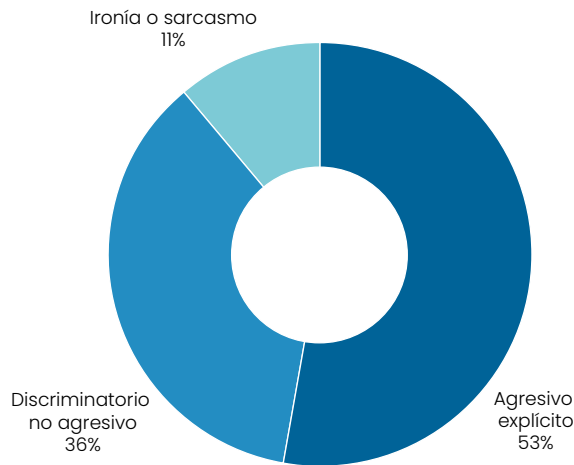
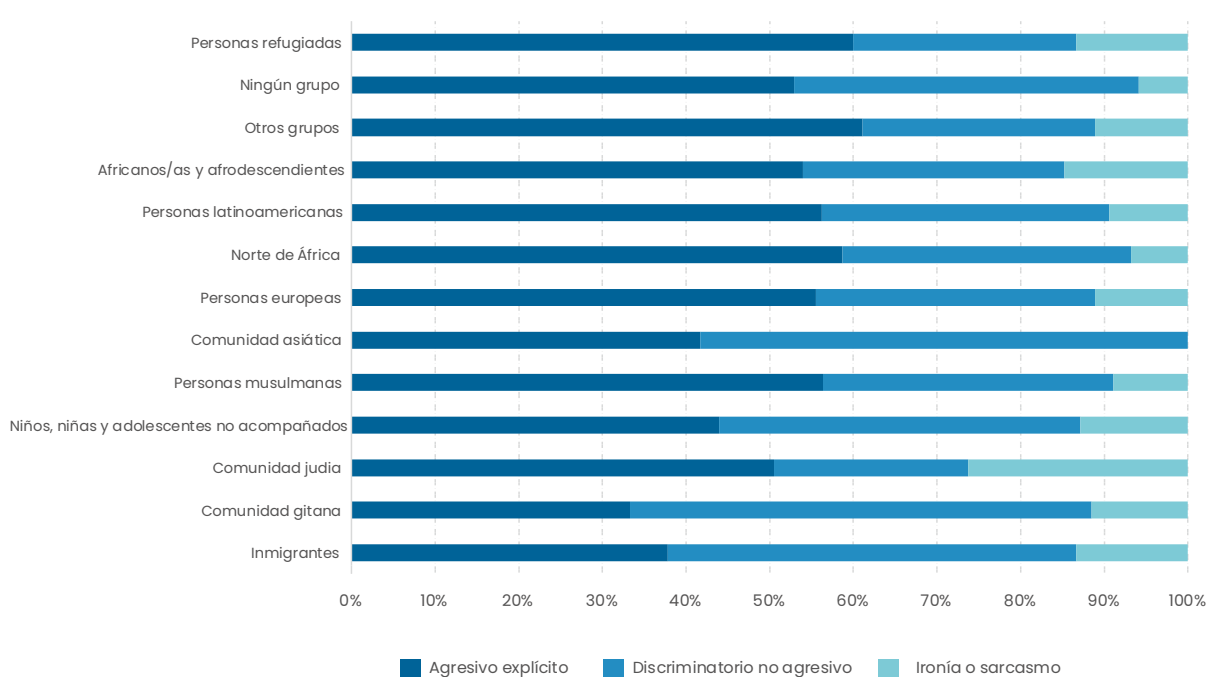
La tipología del discurso de odio en función de la reacción de las plataformas de retirarlo se muestra en el gráfico 8. Se observa que la mayoría de las publicaciones que incitan a la violencia con amenazas directas o indirectas (57%) no fueron retiradas, mientras que solo un 6% fueron eliminadas dentro de las primeras 24 horas. De manera similar, el contenido que deshumaniza o degrada gravemente a los grupos en un 59% no se ha retirado, y solo un 4% fue retirado en el primer día. Cabe destacar que las publicaciones que alaban a quienes atacan al grupo target presentan una tasa de no retirada del 55%, y un 33% de estas fueron a través de la *vía trusted flagger*. En contraste, los contenidos que incitan a la expulsión del colectivo y los que promueven el descrédito en base a atributos personales mostraron una mayor proporción de no retirada, con 70% y 72% respectivamente, reflejando una menor acción inmediata frente a estas formas de discurso de odio.

4.2.3. Expresión del discurso de odio

De las tres categorías que se han considerado para la expresión del discurso de odio (gráfico 9), los resultados muestran que el **discurso agresivo explícito** es el más frecuente, ya que se observa en el 53% de los contenidos

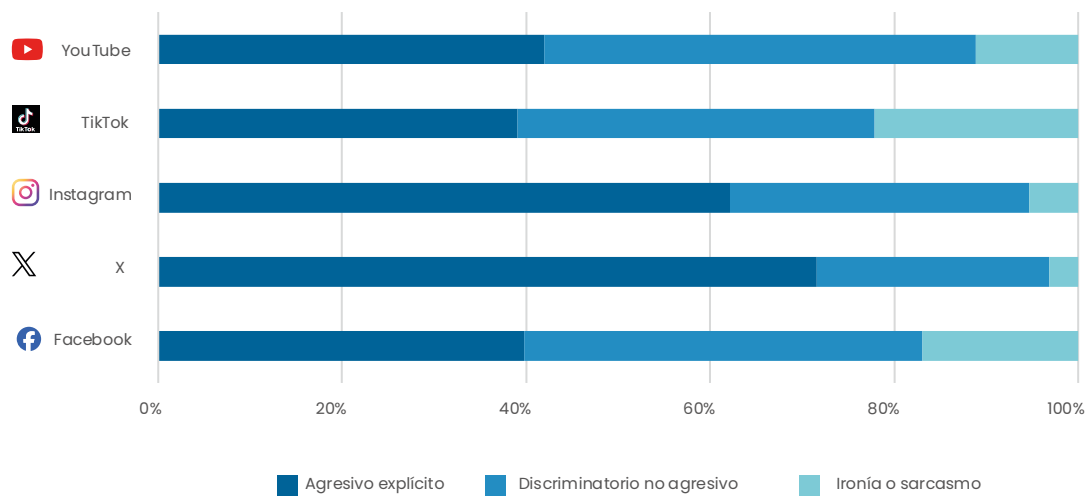
notificados. Ejemplos de contenidos monitorizados: *"La única forma es ir todos a una y empezar una caza contra estas ratas..."*; *"Esas basuras hay que eliminarlas"*. El **discurso discriminatorio no agresivo** se observa en el 36% de los contenidos monitorizados con ejemplos como los siguientes: *"Hay que echarlos a todos los extranjeros si no esto se va de las manos"*; o *"Inmigración... No falla"*. Y el **tono irónico o sarcástico** está presente en el 11% de los contenidos. Algunos de los contenidos monitorizados son: *"otro mono que se escapó del zoo"*; *"si estuviera el pintor austriaco ibais a entrar en las chimeneas"*. La ironía y el sarcasmo ha aumentado en el último año cuatro puntos porcentuales. Este incremento evidencia una creciente complejidad en las estrategias comunicativas empleadas para difundir mensajes discriminatorios. Estas expresiones, disfrazadas de ambigüedad y codificadas culturalmente como humor o crítica, dificultan su reconocimiento social como discursos de odio, lo que favorece su normalización y reproducción en los espacios digitales.

En referencia a la expresión del discurso de odio empleado según el grupo diana al que va dirigido, se observan algunas variaciones, aunque en la mayoría de los casos sigue predominando el uso de un discurso agresivo explícito. Este tipo de discurso se presenta en un 61% de los casos dirigidos a otros grupos, en un 60% hacia las personas refugiadas, en un 59% hacia personas del norte de África y en un 56% en el contexto de la islamofobia. No obstante, el gráfico 10 revela que en el caso del discurso de odio dirigido hacia la comunidad asiática (58%), la comunidad gitana (55%) y las personas inmigrantes (49%), se observa una mayor prevalencia del discurso discriminatorio no agresivo.

Gráfico 9. Frecuencia de la expresión del discurso de odio**Gráfico 10. Distribución de la expresión del discurso de odio según el grupo diana al que se dirige**

El gráfico 11 muestra la distribución del tipo de discurso de odio en función de su presencia en cada una de las plataformas. Se observa que en X e Instagram predomina el discurso agresivo explícito, con un 72% y un 62% de los contenidos notificados a estas plataformas respectivamente. Por el contrario, el discurso discriminatorio no agresivo es más frecuente, en YouTube (47) y en Facebook (43%). En cambio, el tono irónico se manifiesta con mayor frecuencia en TikTok, donde alcanza el 22% de los casos notificados.

La complejidad para identificar el discurso de odio cuando se utiliza un tono irónico o sarcástico queda evidenciada en los datos sobre retirada de contenidos por parte de las plataformas. De las 2.870 notificaciones realizadas, 323 corresponden a contenido irónico, de los cuales el 44% fue retirado. En comparación, los contenidos con discurso agresivo fueron eliminados en un 38% de los casos, mientras que los contenidos no agresivos tuvieron una tasa de retirada del 29%.

Gráfico 11. Distribución del tipo de expresión del discurso de odio según la plataforma de internet

4.2.4. Episodio prototípico al que se vincula el discurso de odio

La hostilidad en redes sociales hacia las personas de origen extranjero ha dejado de ser un fenómeno puntual, vinculado exclusivamente a hechos coyunturales. En 2024, esta tendencia se ha consolidado como una manifestación persistente, articulada en torno a estereotipos, bulos y polarización que encuentran en las redes sociales una vía rápida para su propagación. Aunque algunos picos de actividad responden a acontecimientos concretos, 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.

El 34% de los comentarios y/o imágenes racistas y xenófobas están relacionados con la atribución a las personas de origen extranjero de provocar violencia y/o robos. La vinculación con la inseguridad ciudadana sigue latente en 2024, siendo el incidente más predominante en los contenidos comunicados a las plataformas. Una parte considerable de estos contenidos promueve la idea de que la inmigración es sinónimo de amenaza, apelando a la necesidad de "fronteras seguras" frente a una supuesta "invasión migratoria" y defendiendo incluso la deportación masiva de determinados grupos diana, especialmente de las personas del norte de África.

Cabe señalar 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; por lo que pueden hacer referencia a información falsa, bulos o incidentes descontextualizados.

Un ejemplo de ello fueron los bulos y/o falsas noticias que se difundieron rápidamente en las redes sociales, con el objetivo de atribuir la inseguridad ciudadana a las personas de origen extranjero. Un caso destacado fue el relacionado con el asesinato de un niño en Mocejón (Toledo) ocurrido en agosto, que generó una alarma social y promovió falsamente la percepción de las personas inmigrantes, especialmente las del norte de África, como una amenaza. Este episodio, además, contribuyó a la fragmentación social, en un contexto marcado por el debate sobre las políticas migratorias y la gestión de los centros de acogida de menores extranjeros.

Otro ejemplo importante ha sido el de la catástrofe de la DANA ocurrida el 29 de octubre, que afectó de manera particular a la provincia de Valencia, pero también a Castilla-La Mancha y Andalucía. En este caso, se criminalizó a diferentes grupos diana, como las personas del norte de África, las personas musulmanas y la comunidad gitana, relacionándolos con saqueos y robos. Este enfoque estigmatizó a estos grupos como amenazas para la ciudadanía en un momento de duelo y vulnerabilidad, incertidumbre y desgaste emocional para la población de las zonas afectadas. Posteriormente, en las semanas marcadas por la solidaridad ciudadana, se extendió una narrativa de discurso de odio, basada en bulos, que difundió que las personas de origen inmigrante, y particularmente las mujeres musulmanas, no estaban prestando ninguna ayuda en los pueblos afectados; y, sin embargo, sí se aprovechan, supuestamente, de las subvenciones del Estado y de los impuestos de la ciudadanía española, alimentando la hostilidad hacia las personas inmigrantes.

Pero no solo los acontecimientos nacionales activan estos discursos, sino que también hechos ocurridos fuera de España, como el "atentado terrorista" ocurrido en Mag-

deburgo (Alemania) suscitan comentarios de discurso de odio, los cuales se vinculan con la percepción de inseguridad ciudadana. Estos mensajes buscan generalizar el miedo y la desconfianza hacia determinados grupos, promoviendo el miedo en la población.

Sin embargo, una proporción significativa del discurso de odio monitorizado, un 21%, no responde a **ningún episodio prototípico concreto**. Esto indica que la hostilidad hacia las personas migrantes se produce de manera autónoma, al margen de los acontecimientos sociales específicos. Esto sugiere una base estructural de prejuicios y actitudes discriminatorias que se mantienen activas de forma constante, alimentadas por desinformación, rumores y contenido conspirativo. Lo que indica que el discurso anti inmigratorio que se difunde a través de las redes sociales, principalmente a través de *fake news*, también está suscitado por los estereotipos hacia las personas de origen extranjero.

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. El 9% del total de los contenidos monitorizados en 2024 están vinculados directamente con este episodio prototípico. En febrero, coincidiendo con la noticia del traslado de más de mil inmigrantes desde Canarias a la península, se produjo un pico del 19% de los casos, con frases como *"Lamentable noticia pues solo traen enfermedades y desgracias"*. Otro pico se observó en abril, cuando nuevos traslados generaron un 16% de los discursos de odio relacionados con políticas migratorias. La tendencia muestra una correlación entre decisiones de las administraciones públicas orientadas a la inclusión y respuestas sociales marcadas por la hostilidad y la desinformación.

El discurso ligado a "políticas públicas" muestra una tendencia variable. Hay una fluctuación a lo largo de los meses, con periodos de baja y picos esporádicos que reflejan posiblemente el discurso racista asociado a la actividad en centros de acogida de personas inmigrantes, así como a la gestión en política migratoria y la situación en Canarias respecto a la acogida de niños, niñas y jóvenes no acompañados; junto con el traslado a centros de otras comunidades autónomas.

A esta narrativa se suman los contenidos vinculados a la **llegada de embarcaciones a las costas** canarias; con especial intensidad a partir de junio, y que generaron una nueva ola de discurso de odio. El 7% de los contenidos registrados en todo el año estuvieron relacionados con este episodio, concentrando un 14% en junio tras la noticia del rescate de 10 embarcaciones con más de 500 personas a bordo. Los mensajes más repetidos fueron de violencia simbólica extrema, con expresiones como "¿Y la Armada para qué quiere los

misiles?" o *"Cemento y al fondo del mar"*. Este tipo de comentarios, además de banalizar la muerte y el sufrimiento humano, normalizan discursos de exterminio que, aunque ilegales, algunos circulan con impunidad en las plataformas digitales.

Asimismo, el ámbito deportivo, que representa el 4% de los contenidos monitorizados en 2024, ha sido también un escenario donde se ha expresado de manera intensa el racismo y la xenofobia. A lo largo del año, se han identificado picos significativos de discurso de odio vinculado a eventos deportivos, con una incidencia especialmente alta en el fútbol. En este contexto, el racismo ha sido dirigido principalmente a jugadores, como en el caso de Vinicius Júnior, cuyo episodio de discriminación volvió a poner en el centro del debate el racismo tanto en los estadios como en las plataformas digitales. Este fenómeno alcanzó su punto álgido en marzo, con un 18% de los contenidos relacionados con este caso específico. Durante este periodo, los insultos, las burlas y los cuestionamientos identitarios fueron recurrentes, como se ejemplifica con expresiones como *"Monicius en estado puro 🤔🤔🤔"*.

El impacto del discurso de odio en el ámbito deportivo no se limita únicamente a los comentarios en redes sociales. También se han producido incidentes violentos y manifestaciones intolerantes en los estadios. Un ejemplo claro ocurrió el 7 de noviembre, durante un partido de la Europa League entre el Ajax de Ámsterdam y el Maccabi Haifa de Israel, donde se registraron cánticos xenófobos y agresiones físicas entre los aficionados. Los contenidos de odio identificados en este caso se dirigían tanto a personas judías como musulmanas.

Además, el fenómeno del discurso de odio no se circunscribe solo al fútbol. En los Juegos Olímpicos de París en julio y agosto, que coincidieron con la participación de selecciones como la de Marruecos, la narrativa de odio alcanzó una mayor visibilidad. También se produjo un pico relevante en la Supercopa de la UEFA, donde se registró un aumento del 29% en los mensajes de odio enmarcados en lo deportivo. Estos mensajes incluían comentarios discriminatorios hacia atletas de origen extranjero que representaban a España, bajo cuestionamientos identitarios y también hacia un jugador de la selección de Marruecos que recibió una oleada de mensajes racistas como *"Aunque haya nacido en 🇪🇸 no deja de ser un mor. de 🍌"*. Estos comentarios reflejan la hostilidad hacia la identidad de los deportistas, que, a pesar de representar a su país, son percibidos como «extranjeros» por algunos sectores de la sociedad, lo que pone de manifiesto el racismo y la xenofobia subyacentes en muchos de estos discursos.

El discurso de odio en el ámbito deportivo no solo se limita a la verbalización de insultos racistas o xenófobos,

sino que crea barreras que dificultan la integración de las personas de origen extranjero.

También, se registraron importantes picos de contenido islamóforo, coincidiendo con fechas simbólicas del calendario musulmán o tras atentados ocurridos en otros países europeos. Muchos de estos discursos consistieron en la acusación generalizada de que el islam es incompatible con los valores democráticos occidentales. Las mujeres musulmanas, especialmente aquellas que llevan hiyab, fueron las principales víctimas de estos ataques, siendo presentadas como símbolos de una supuesta *"opresión cultural"* o incluso como amenazas a la *"libertad de las mujeres españolas"*. El discurso islamóforo estigmatiza, genera miedo y condiciona la vida diaria de miles de personas que profesan el islam en España.

Asimismo, el antigitanismo se manifiesta a través de vídeos de conflictos vecinales en barrios, que son utilizados para criminalizar colectivamente a la comunidad gitana, reforzando prejuicios sobre su presunta vinculación con la delincuencia, la informalidad económica o la violencia. Este tipo de contenidos, cargados de lenguaje despectivo y humor discriminatorio, siguen alimentando una percepción negativa que impacta directamente en la integración social de la minoría étnica más discriminada en España.

El antisemitismo también resurgió, particularmente a partir del conflicto en Gaza. Se detectaron numerosas publicaciones con mensajes que negaban el Holocausto, caricaturas y simbología nazi. También, algunas publicaciones contenían una narrativa que responsabilizaba colectivamente a la comunidad judía de todos los eventos violentos en el Medio Oriente. Este tipo de discurso no hacía distinción entre la comunidad judía en su conjunto y las acciones del gobierno de Israel, perpetuando la idea de que las personas judías, en su conjunto, son responsables de los conflictos bélicos.

En conjunto, los datos muestran que el discurso de odio no se limita a momentos de crisis o conflicto, sino que encuentra en diversos episodios, sean reales, ficticios o simbólicos, un catalizador para su propagación. Las redes sociales se consolidan como un espacio donde la hostilidad hacia las personas de origen extranjero se naturaliza.

En el gráfico 12 se muestra la proporción de contenidos notificados a las plataformas según el episodio prototípico que los ha desencadenado o al que se vinculan.

El análisis de 2024 confirma que el discurso de odio en redes sociales no es solo el resultado de momentos de crisis, sino un fenómeno estructural. La difusión constante de odio, en forma de comentarios, memes o noticias falsas, no

Gráfico 12. Distribución del discurso de odio notificado a las plataformas según su vinculación con un episodio prototípico

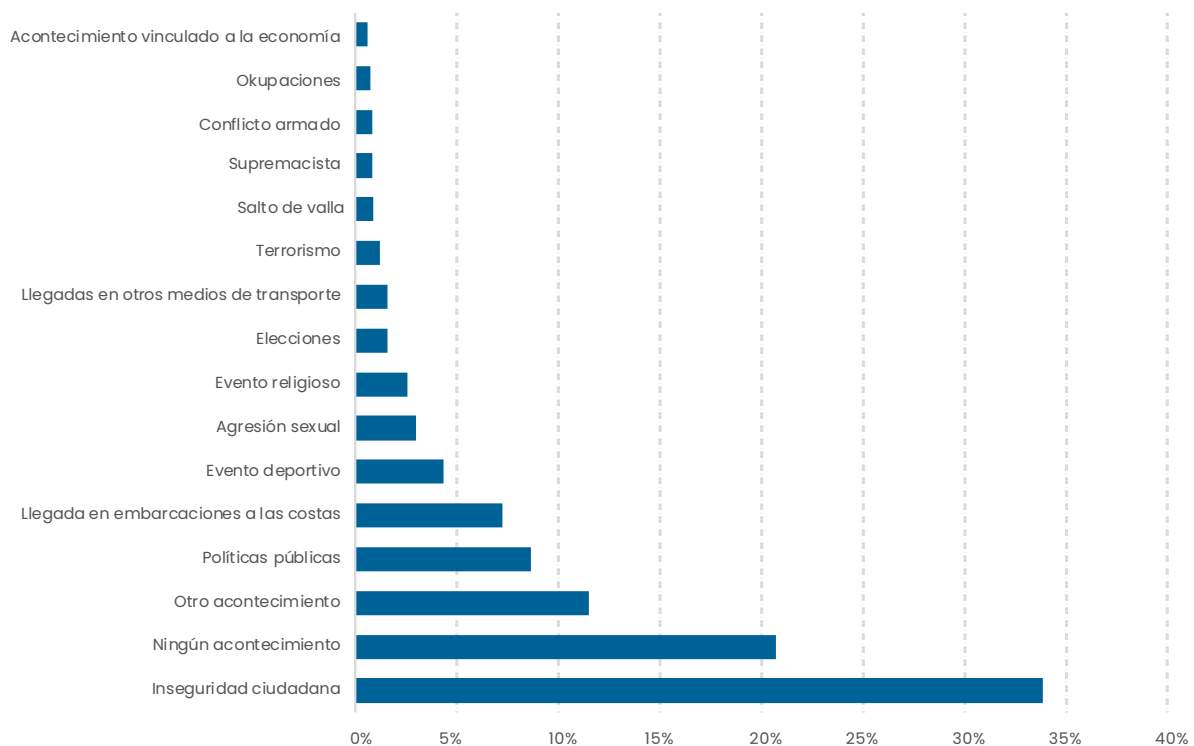
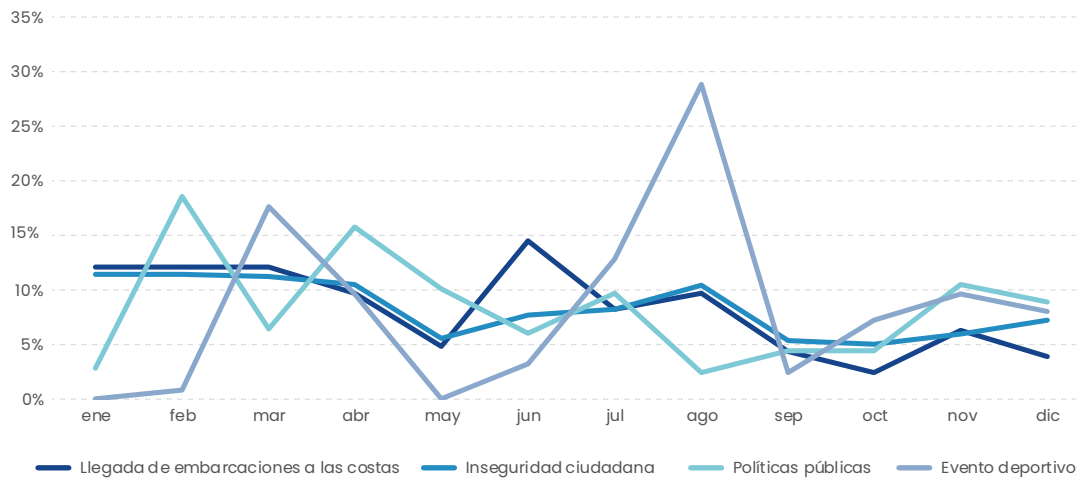


Gráfico 13. Evolución de los principales episodios prototípicos en 2024

solo afecta a las personas migrantes, sino que debilita los principios fundamentales de una sociedad democrática.

A continuación, se presenta la evolución a lo largo del año del discurso de odio a partir de los principales episodios prototípicos identificados durante el año (gráfico 13). Los datos muestran que los repuntes en los discursos hostiles están estrechamente vinculados a eventos coyunturales fácilmente reconocibles.

La llegada de embarcaciones a las costas actuó como un detonante constante, con tasas superiores al 12% en los tres primeros meses del año y un repunte en junio (14%). Esta narrativa refuerza la idea de una amenaza externa y es utilizada de forma recurrente para legitimar discursos de rechazo hacia personas migrantes y refugiadas.

Asimismo, la inseguridad ciudadana fue una narrativa estructural en los discursos de odio, con presencia sostenida y notoria en los primeros meses del año, alcanzando su pico en enero y febrero (ambos con 11%) asociándose con la criminalización de determinados colectivos.

Por su parte, las políticas públicas adquirieron un protagonismo destacado en febrero (19%) y en abril (16%), coincidiendo con debates institucionales que provocaron fuertes reacciones en redes sociales. Estos debates generaron un rechazo marcado hacia las poblaciones percibidas como beneficiarias de ayudas o medidas de inclusión, especialmente migrantes y minorías, convirtiéndose en un foco recurrente de discursos hostiles.

Finalmente, los eventos deportivos se revelaron como un desencadenante poco común, pero de gran impacto, alcanzando un pico notable en agosto (29%) coincidiendo con competiciones internacionales (por ejemplo, los Juegos Olímpicos de París). En esos contextos, se evidenció un aumento de expresiones xenófobas vinculadas al rendimiento de deportistas de origen extranjero o la representación de países determinados, reflejando como el deporte puede funcionar como un factor de identidades nacionales excluyentes.

Por su parte, el gráfico 14 recoge los episodios prototípicos según el grupo de población diana al que afectan más. La inseguridad ciudadana es el episodio prototípico que predomina en todos los grupos diana, aunque es más relevante en los discursos dirigidos hacia las personas del norte de África en los que el 51% de las comunicaciones relativas a este grupo están ligadas a este episodio prototípico. Del mismo modo ocurre con las personas latinoamericanas (48%) y niños, niñas y adolescentes no acompañados (47%).

En cuanto al episodio prototípico de políticas públicas, se observa que un 23% de los casos están relacionados con los niños y niñas menores no acompañados y un 16% con la inmigración.

Por otra parte, el episodio prototípico de llegada de embarcaciones a las costas está vinculado a un 19% de los casos a las personas inmigrantes.

INFORME ANUAL DE MONITORIZACIÓN DEL DISCURSO DE ODIOS EN REDES SOCIALES

Gráfico 14. Distribución de episodios prototípicos relacionados con el discurso de odio según grupo diana

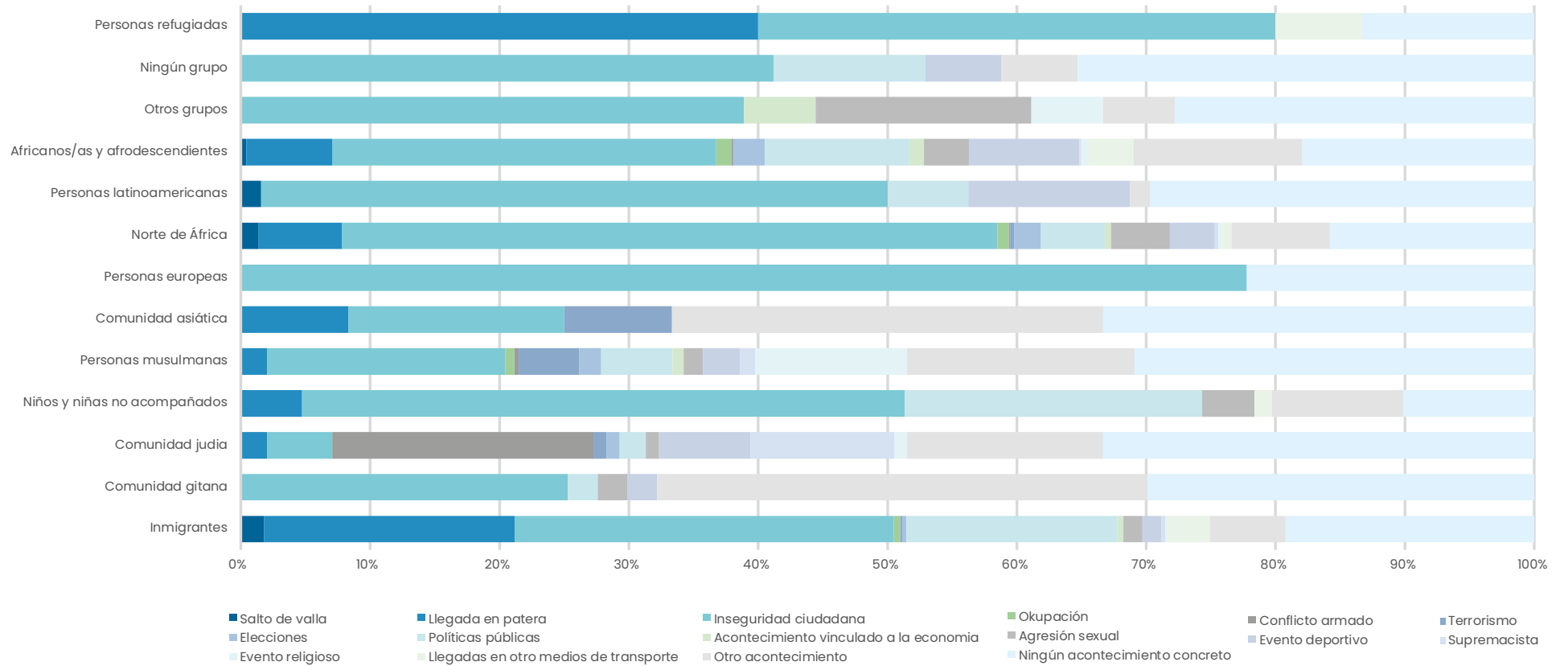
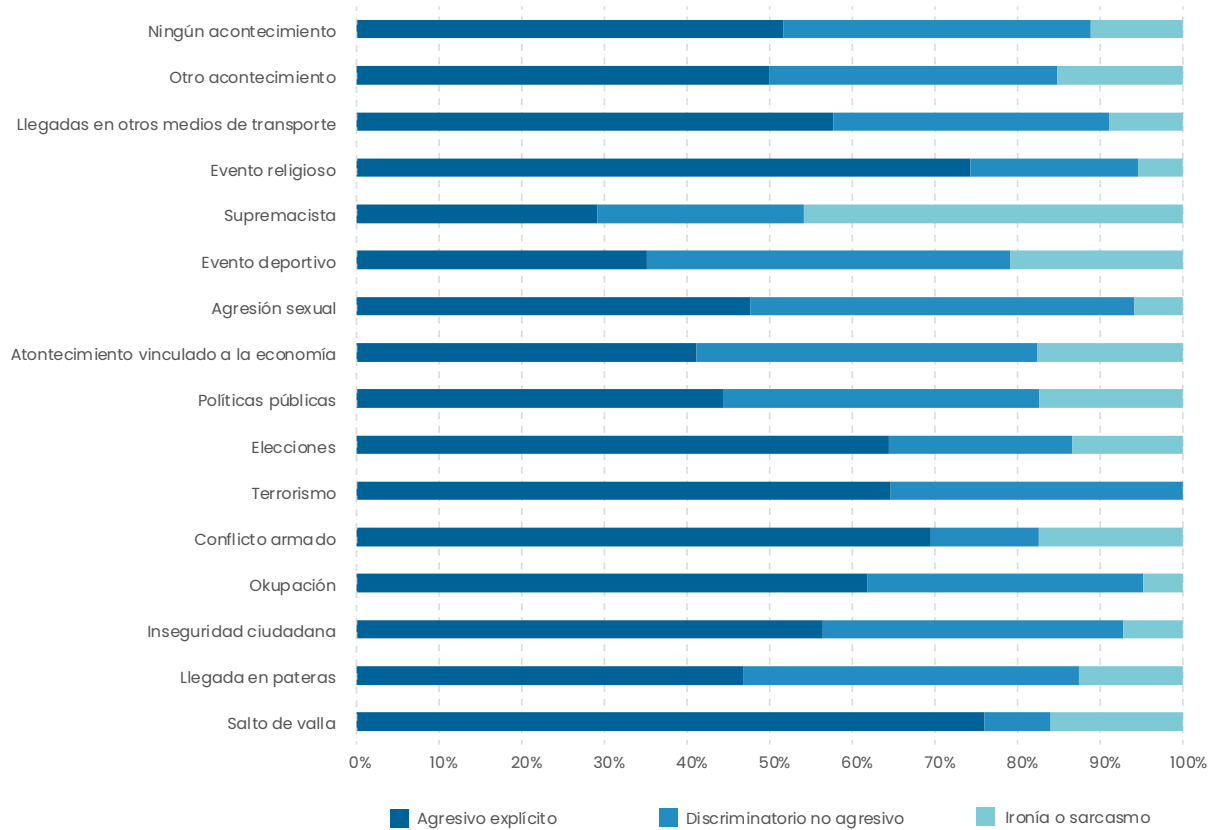


Gráfico 15. Distribución de la expresión del discurso de odio según el episodio prototípico

El gráfico 15 muestra la relación entre los episodios que originan el discurso de odio y la naturaleza del lenguaje utilizado. El episodio más frecuente identificado es el de "inseguridad ciudadana", en el que predomina el uso de un discurso agresivo explícito, presente en el 56% de las comunicaciones vinculadas a este episodio. También, este tipo de lenguaje predomina en situaciones como los saltos a las vallas de Ceuta y Melilla (76%), eventos religiosos (74%) y conflictos armados (70%).

En cuanto a los discursos de odio menos explícitos, el lenguaje discriminatorio no agresivo es el más frecuente en contextos como agresión sexual (46%) y eventos deportivos (44%). Además, los discursos con tono de ironía o sarcasmo se vinculan con episodios de carácter "supremacista", alcanzando un 46% de los contenidos analizados relacionados con este episodio.

4.2.5. El discurso de odio hacia las mujeres

El discurso de odio en las redes sociales, al igual que en otros ámbitos de la sociedad, refleja y refuerza las dinámicas de discriminación, odio e intolerancia que se encuentran presentes en nuestra sociedad.

Las personas y colectivos pueden ser objeto de discurso de odio por diferentes motivos o características personales ("raza", color, origen nacional o étnico, edad, discapacidad, religión, sexo, identidad de género y orientación sexual) o combinaciones de estos. A este respecto, a la hora de hacer seguimiento del discurso de odio y de analizar sus tendencias, el género es una de las dimensiones más importantes a tener en consideración.

El análisis de los datos monitorizados en relación con el discurso de odio en 2024 muestra cómo los comentarios sexistas, racistas, xenófobos e islamófobos van a menudo de la mano, afectando de manera particular a ciertos grupos de mujeres. De acuerdo con los datos, aunque la mayoría de los contenidos analizados están relacionados con el masculino genérico (90%), un 6% de los contenidos presentan una marca lingüística en femenino. Esta cifra, aunque aparentemente baja, revela una presencia de expresiones o manifestaciones de odio dirigidas específicamente a las mujeres, y particularmente a determinados grupos diana.

Así, un alto porcentaje de los contenidos con marca lingüística en femenino se dirigen hacia las mujeres musulmanas (60%) y afrodescendientes (21%).

En el caso de las mujeres musulmanas, en particular, se observa una doble o triple discriminación: por su género, su origen y/o pertenencia religiosa. Los comentarios hacia estas mujeres, los términos peyorativos con los que se las describe están, con frecuencia, asociados con estereotipos negativos sobre el islam, y basados en estereotipos de género arraigados, percibiéndolas como sumisas o subordinadas. Ejemplos como *"Está enferma mental, vamos a ver cuánto le dura la risa si se casa con un musulmán y sus costumbres"* o *"Las sombrillas son para el verano... menuda ruina pestosa"* ponen de manifiesto una doble o triple discriminación hacia estas mujeres.

Un elemento recurrente en este tipo de discurso es la referencia al uso del hiyab o velo islámico, un símbolo que, en ocasiones, se convierte en objeto de crítica y burla. El rechazo al hiyab no solo se basa en prejuicios hacia la religión musulmana, sino que, lejos de ser comprendido como una elección personal o un símbolo religioso, su uso es presentado como una forma de opresión, alimentando así un discurso que niega la autonomía de las mujeres musulmanas. Este rechazo no se limita a una crítica religiosa, sino que forma parte de un discurso más amplio que cuestiona su capacidad de decidir por sí mismas y refuerza estereotipos sexistas, racistas e islamófobos.

Del mismo modo, los comentarios hacia las mujeres afrodescendientes demuestran una vez más la intersección de la misoginia y el racismo. Expresiones como *"No necesito que esa negra de mierda me defienda"* o *"En Asturias cultivan algodón y alguien lo tiene que recoger"* ejemplifican cómo el odio hacia las mujeres afrodescendientes se construye sobre una narrativa racista y sexista, reduciéndolas a roles tradicionalmente estereotipados. Este tipo de comentarios no solo evidencia el racismo, sino también el sexismo persistente que afecta a las mujeres.

Al analizar la relación de los contenidos con los episodios prototípicos que suscitan este tipo de discursos, encontramos que un 32% de los comentarios con marca lingüística femenina no están vinculados a ningún acontecimiento específico. Este dato es particularmente relevante, ya que refleja que el odio hacia las mujeres se reproduce de forma constante, alimentándose de actitudes, estereotipos y prejuicios sociales arraigados. No obstante, también se observa que un 13% de los contenidos están relacionados con la inseguridad ciudadana, lo que pone de manifiesto cómo el miedo colectivo ante personas o grupos que son percibidos como diferentes o la violencia social se proyecta sobre las mujeres, percibiéndolas como una amenaza para la convivencia social o la identidad cultural del grupo mayoritario.

Comentarios como *"Las mujeres estamos amenazadas por la Chusma enemigos del jamón"* o *"Españistan. La dictadura feminista terrorista musulmana de la violencia. La maldad y la muerte"* son ejemplos de cómo el discurso de odio utiliza el miedo al diferente y a lo "no occidental" para justificar la violencia hacia las mujeres. Estos comentarios incitan a la xenofobia y el racismo, y también alimentan la preocupación social por los cambios culturales y sociales, presentando a dichas mujeres como una amenaza para la ciudadanía.

En este contexto, la mirada interseccional resulta indispensable para comprender la complejidad del discurso de odio hacia estos grupos, y cómo este se articula en múltiples capas que combinan sexismo, racismo, islamofobia y xenofobia. Las mujeres de origen extranjero, mujeres afrodescendientes o mujeres musulmanas en España no solo reciben odio por su género, sino también por su pertenencia a grupos socialmente estigmatizados, lo que las posiciona en una situación especialmente vulnerable.



5. CONCLUSIONES

El análisis del discurso de odio en redes sociales a lo largo de 2024 se muestra como un fenómeno estructural y persistente, que trasciende los hechos coyunturales para convertirse en un elemento constante en las plataformas digitales. A pesar de que ciertos picos de actividad se vinculan a acontecimientos específicos como traslados de personas migrantes, catástrofes naturales o eventos deportivos, una proporción significativa de los mensajes hostiles persiste al margen de cualquier suceso concreto. Esta continuidad pone en evidencia una base sólida de racismo, xenofobia, islamofobia, antigitanismo y antisemitismo, que encuentra en las redes sociales un espacio para su reproducción y normalización.

Los datos obtenidos muestran que las cinco plataformas monitorizadas (X, Facebook, Instagram, TikTok y YouTube) retiraron tan solo un 36% de los contenidos que se les notificaron en 2024, que podían ser constitutivos de delito, de infracción administrativa o que violaban sus propias normas de conducta en internet.

En cuanto a los grupos diana, las personas originarias del norte de África, seguidas de las personas afrodescendientes, inmigrantes y personas musulmanas son las cuatro categorías diana a las que más contenido de odio se dirige (35%, 24%, 21% y 20% respectivamente sobre el total monitorizado). También, aunque en menor

porcentaje, es reseñable que un 5% de los contenidos se dirige hacia los niños, niñas y jóvenes no acompañados.

Uno de los principales datos clave del informe es que la inseguridad ciudadana, con un 34% de los contenidos reportados, sigue siendo el episodio prototípico más vinculado a los discursos de odio, especialmente hacia las personas del norte de África y niños, niñas y adolescentes no acompañados. Esta narrativa, que asocia inmigración con delincuencia, ha sido alimentada por la desinformación, los bulos y noticias manipuladas, que en su mayoría no responden a hechos reales ni recientes, y que en un 40% de los casos consisten en contenido falso o descontextualizado. La constante apelación al miedo colectivo, sostenida por estereotipos de peligrosidad, facilita la justificación de medidas represivas como la deportación o el cierre de fronteras, socavando principios democráticos y derechos fundamentales de las personas.

Además, la instrumentalización de episodios concretos como el asesinato en Mocejón o la catástrofe de la DANA en Valencia demuestra cómo determinadas crisis pueden ser aprovechadas para alimentar el discurso de odio. En estos casos, no solo se criminaliza a las personas migrantes, sino que se alimenta una percepción de amenaza existencial que fragmenta la cohesión social y genera climas de polarización e intolerancia. La narrativa que cuestiona la legitimidad de las personas

extranjeras para recibir ayudas o participar en procesos de solidaridad acentúa aún más esta exclusión simbólica, presentándolas como beneficiarias inmerecidas y no como ciudadanos/as de pleno derecho.

Asimismo, se constata que los discursos de odio no se limitan a los acontecimientos del territorio nacional. Incidentes internacionales, también desencadenan oleadas de comentarios hostiles en redes sociales en España. Este fenómeno demuestra la existencia de un imaginario globalizado de la amenaza migrante o islámica, donde los grupos estigmatizados son tratados como un bloque homogéneo, y cualquier incidente es utilizado como justificación para reforzar el rechazo colectivo.

Otro de los episodios prototípicos que generan el discurso de odio es el vinculado a las políticas públicas (9%). Las políticas públicas, como el traslado de personas desde Canarias a la península, provocan reacciones negativas que expresan una fuerte oposición a la inclusión de personas migrantes en el tejido social. Estas narrativas, que denuncian supuestos privilegios institucionales, generan una percepción de injusticia entre la población autóctona y erosionan la confianza en las instituciones democráticas. Este tipo de discursos ha mostrado un patrón fluctuante, con picos que coinciden con decisiones administrativas específicas, lo cual revela una sensibilidad social ante los temas migratorios y la gestión de la diversidad.

El ámbito deportivo, con un 4% de los contenidos reportados también se configura como un espacio donde se reproducen y amplifican discursos de odio, particularmente racistas y xenófobos. Casos de jugadores de fútbol ponen de relieve cómo el racismo se expresa de forma abierta tanto en los estadios como en las redes sociales. La hostilidad dirigida hacia atletas de origen extranjero, incluso cuando representan a España en competiciones internacionales, revela un conflicto profundo sobre la identidad nacional y la pertenencia. Este tipo de discursos no solo afectan a los deportistas individuales, sino que tiene un efecto simbólico de exclusión hacia comunidades enteras, al negar su integración plena en la sociedad.

Especial mención merece el impacto del discurso de odio dirigido a mujeres (5%), particularmente aquellas que se encuentran en la intersección de múltiples formas de discriminación, como las mujeres musulmanas y afrodescendientes. La combinación de sexismo, racismo e islamofobia da lugar a una violencia simbólica específica que refuerza estereotipos coloniales y patriarcales, y que se expresa a través de burlas, desprecios o cuestionamientos a su capacidad de decidir sobre su propia vida. Las mujeres que usan hiyab, por ejemplo, son objeto recurrente de ataques que no solo niegan su autonomía, sino que también las presentan como amenazas culturales.

Todo ello refuerza la necesidad de aplicar una mirada interseccional al análisis del discurso de odio, entendiendo que los distintos ejes de discriminación (raza, género, religión, clase) no actúan de forma aislada, sino que se entrelazan y refuerzan mutuamente, generando situaciones de vulnerabilidad extrema.

En cuanto al lenguaje empleado, el análisis revela que los discursos más explícitos y agresivos predominan en contextos como la inseguridad ciudadana, los saltos a las vallas y los conflictos religiosos o armados. Por el contrario, en episodios como los eventos deportivos o las agresiones sexuales, el lenguaje tiende a ser más sutil, a menudo irónico o sarcástico. Esta variabilidad en el tono demuestra que el discurso de odio no siempre se presenta en formas evidentes o denunciabiles, lo que dificulta su detección y sanción por parte de las plataformas digitales. Sin embargo, incluso los comentarios menos explícitos tienen un fuerte impacto simbólico, ya que contribuyen a naturalizar los prejuicios y a crear un clima social hostil hacia determinados grupos.

En cuanto a la tipología del contenido, se constata un patrón de mensajes y/o imágenes en los que se deshumaniza o degrada gravemente a las personas a las que va dirigido (39%), incitando a la violencia con amenazas directas o indirectas (29%). Además se incita a la expulsión del colectivo inmigrante en un 15% de los contenidos monitorizados. Esta estigmatización genera un ambiente hostil. Asimismo, puede alimentar el miedo y el resentimiento hacia ciertos grupos de la población, que pueden generar una mayor fragmentación y conflicto social.

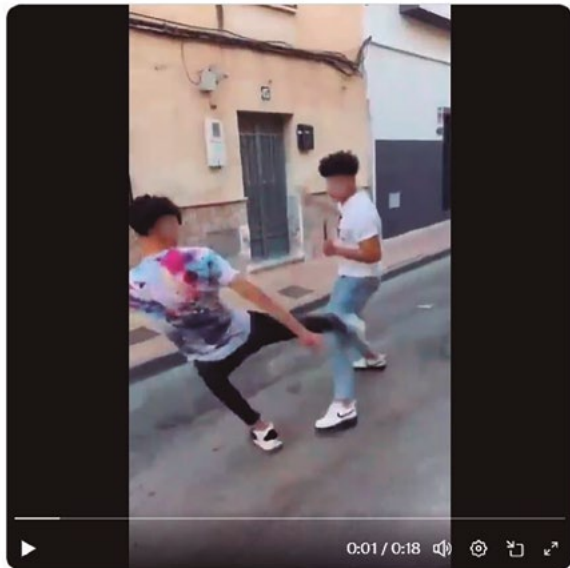
En conjunto, los datos muestran que el discurso de odio en redes sociales en España no es un fenómeno puntual. Se trata de una forma de violencia simbólica estructural que se sostiene en la reproducción de estereotipos, la desinformación y la polarización política y social. El impacto de esta discriminación e intolerancia va más allá de los prejuicios individuales que puede causar a las personas directamente afectadas. Tiene consecuencias graves para la cohesión social, el funcionamiento democrático y los derechos humanos, al erosionar la confianza mutua, alimentar el miedo al otro y debilitar los valores de pluralismo e inclusión.

Es necesario, por tanto, abordar el discurso de odio dirigido hacia las personas de origen extranjero como un fenómeno multifacético que requiere respuestas coordinadas desde las instituciones públicas, las plataformas digitales, los medios de comunicación y la sociedad civil. Las acciones deben ir más allá de la moderación de contenidos y se tienen que realizar acciones de sensibilización y de promoción de narrativas alternativas que contrarresten la deshumanización hacia los grupos diana para construir una sociedad más inclusiva y resiliente frente al odio.

6. ANEXO I: Ejemplos de discurso de odio

A continuación, se presenta una serie de contenidos de discurso de odio en línea monitorizados a lo largo del año 2024 que complementan a los recogidos a lo largo del texto.

1. Episodio prototípico: inseguridad ciudadana y políticas públicas



* El contenido utiliza un video de violencia callejera para atribuir, sin fundamento, conductas delictivas a grupos diana y los presenta como una amenaza social y económica.

2. Discurso de odio contra personas africanas y afrodescendientes

Su comportamiento normal
En su hábitat natural

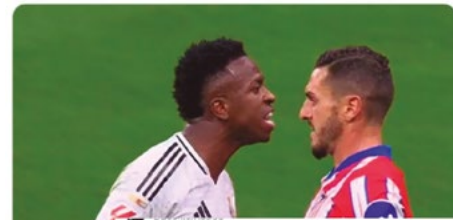


3. Discurso de odio contra personas del norte de África

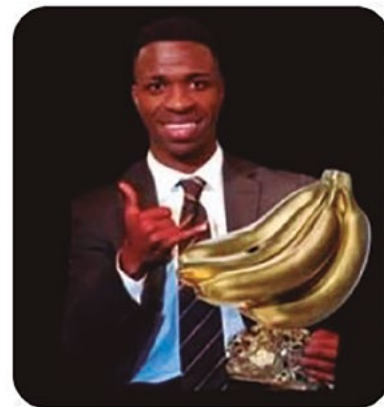
Moro de mierda a ver si te dan una paliza y te dejan seco

4. Episodio prototípico: evento deportivo

No me quiero ni imaginar el POV de Koke.



Desde la perspectiva de Koke.



*Deshumanización hacia el futbolista Vinicius Jr en varios partidos de La Liga de Fútbol en 2024

5. Episodio prototípico: evento religioso



6. Episodio prototípico: salto de valla

Todo esto se arreglaba electrificando la valla,
, ya los demás no lo intentaría

7. Discurso de odio contra niños, niñas y adolescentes menores no acompañados

Illegales menores dilincuentes que es lo que hacen aquí, hacen lo mismo en su país de la esquina ... deportación ya,

8. Discurso de odio contra personas latinoamericanas

Eres un cancer a erradicar, panchito de mierda

9. Mensaje construido con emojis/ lenguaje codificado



*Se utiliza el emojis del ratón para deshumanizar al grupo diana.

Hace falta soltar muchas

OS 🖱️ AS

a mano abierta a todos estos parásitos

🖱️ OP 🖱️ AS.

Deseandito que se lie gordá para hacer PURGA

1.397 Visualizaciones

*El mensaje constituye una incitación a la violencia contra el colectivo. Se combinan emojis con letras para formular mensajes violentos.

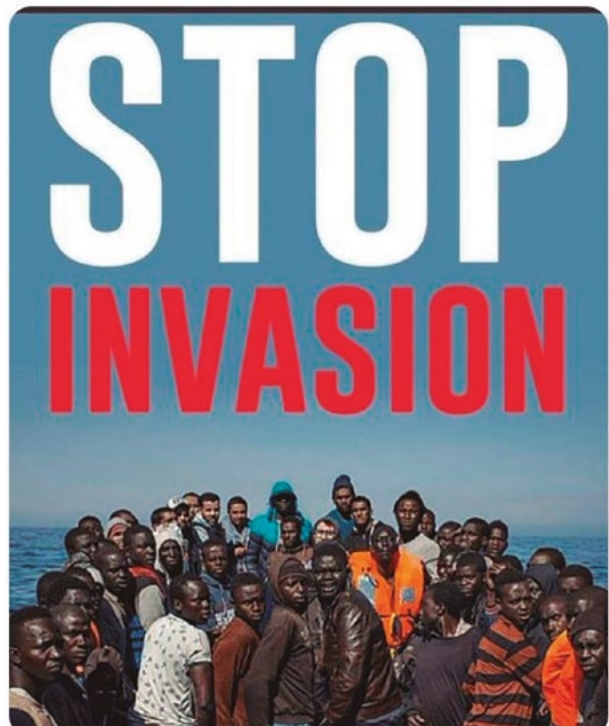
10. Antisemitismo

¡RATAS NARIGUDAS A LA CARRERA!

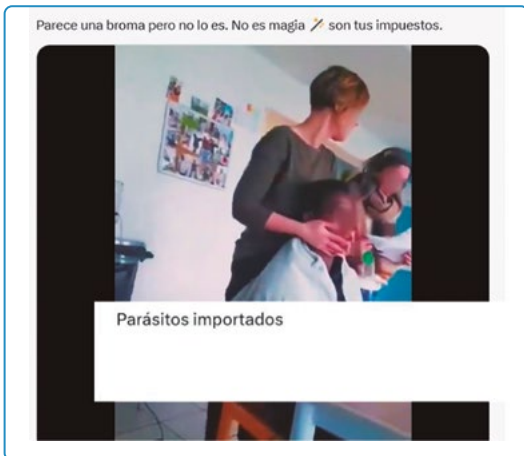


*Los términos "ratas" y "narigudos" se emplean como referencia al Holocausto; como forma de deshumanizar y degradar a las personas judías. Esta imagen está relacionada con el acontecimiento del túnel de Nueva York

11. Episodio prototípico: llegada de pateras



12. Episodio prototípico: políticas públicas



*Esta publicación recurre a la desinformación para generar rechazo hacia menores migrantes no acompañados, presentándolos como beneficiarios injustos de recursos públicos.



13. Se promueve el descrédito en base a atributos personales del grupo



*Se hace uso de una narrativa en la que se presentan a los niños, niñas y adolescentes no acompañados, de origen marroquí, como una amenaza. Para ello se utiliza el “cuchillo” como medio para indicar inseguridad y la “bandera de marruecos” para indicar el origen.



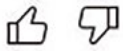
*El término PeloBrócoli es atribuido a las personas del norte de África.

14. Se incita a la expulsión del colectivo

El Mohamed este, de mierda, a su país

15. Se alaba a quien atenta contra el grupo target

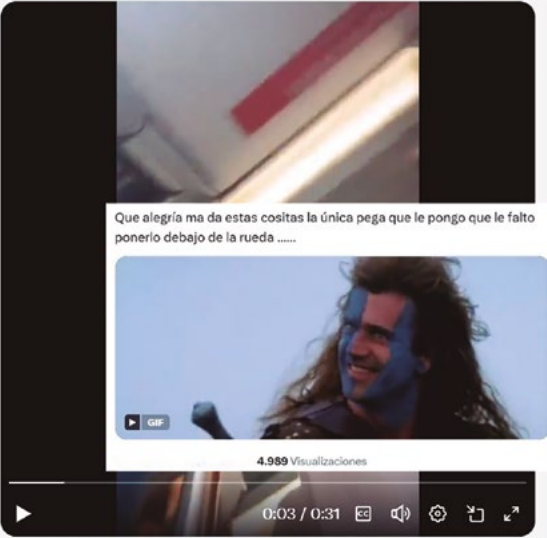
El pintor austriaco tenia razon!



Responder

*El pintor austriaco hace referencia a Hitler

Jovelandeses molestan a las personas que están en el autobús, pero un se relaja 🚩



4.989 Visualizaciones

182 713 4 mil 313 mil

* Esta publicación celebra que un joven increpe a otro identificado como extranjero en un autobús, insinuando que "se relaja" tras ser enfrentado. Además, acompaña el mensaje con una imagen que sugiere que la única falta fue no haberlo "puesto debajo de la rueda", lo que normaliza y glorifica la violencia hacia personas migrantes

16. Se utiliza la ironía o el sarcasmo

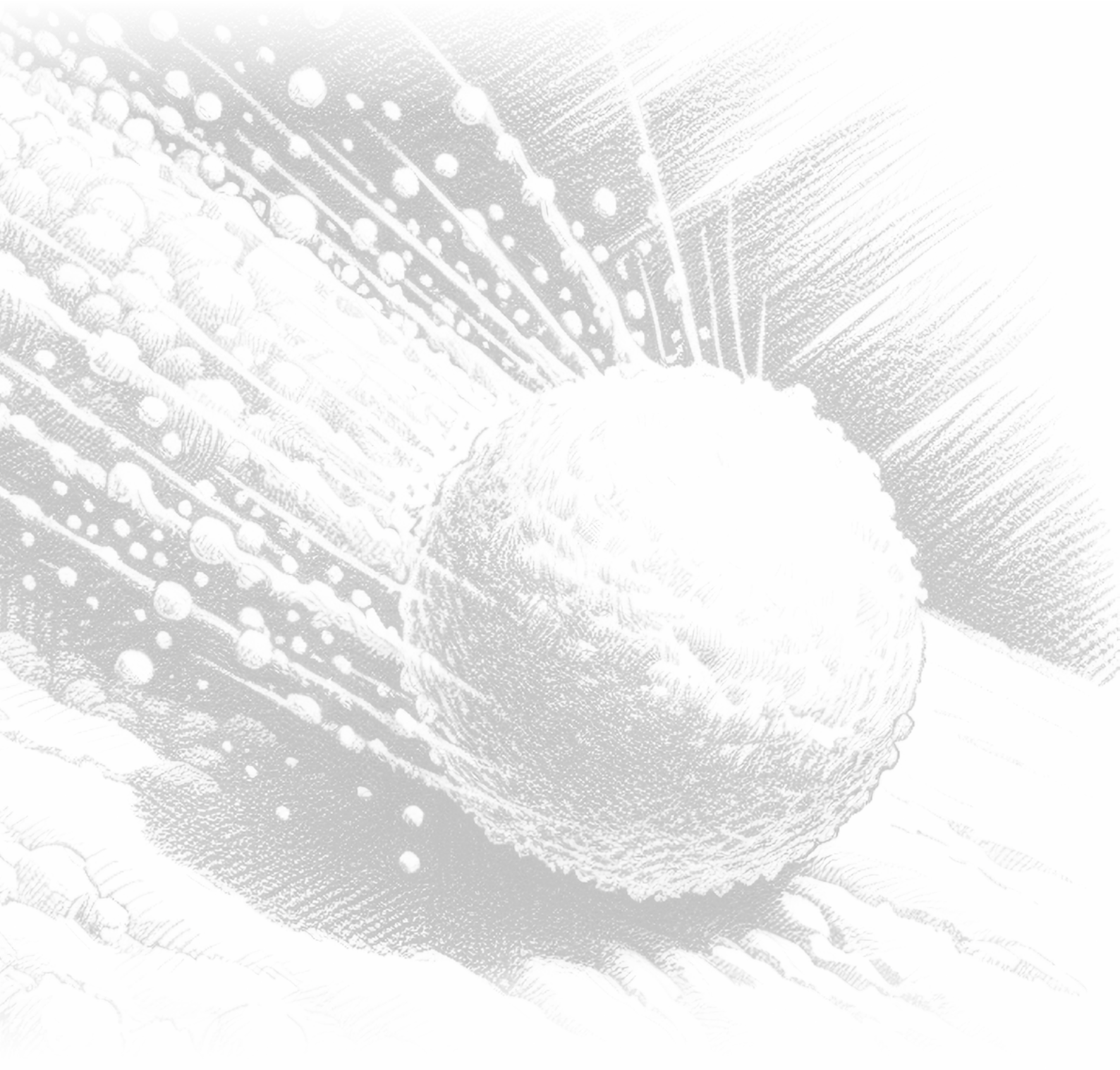
Esto ha pasado en Ferrol, A Coruña:



781 1 mil 4 mil 941 mil

¡Pobres criaturas! Los niños indefensos de los progres, lástima que ninguno los lleva a sus casas.

*Niños indefensos = menores extranjeros



MINISTERIO
DE INCLUSIÓN, SEGURIDAD SOCIAL
Y MIGRACIONES

SECRETARÍA DE ESTADO
DE MIGRACIONES

