ENTREGABLE D2.2

Informe de Base de Datos Centralizada

CIFAL Málaga

Número de Proyecto: 101213796
Convocatoria: CERV-2024-CHAR-LITI
Autoridad concedente: Agencia Ejecutiva Europea de Educación y Cultura
Duración del Proyecto: 24 meses

Cofinanciado por la Unión Europea

Financiado por la Unión Europea. Sin embargo, las opiniones y puntos de vista expresados son exclusivamente los del autor o los autores y no reflejan necesariamente los de la Unión Europea ni los de EACE. Ni la Unión Europea ni la autoridad que concede el préstamo se responsabilizan de ellos.

Detalles contractuales del proyecto

Título del proyecto	Red de Tolerancia contra los delitos de odio
Acrónimo del proyecto	ReTO
N.º del acuerdo de subvención	101213796
Fecha de inicio del proyecto	01.06.2025
Fecha de finalización del proyecto	31.05.2027
Duración	24 meses
Webside	El proyecto no tendrá una página web específica, pero tiene una subpágina en: https://cifalmalaga.org/proyecto/reto/

Detalles del Plan

Paquete de trabajo No.	WP2	Tarea No.	T1.2
Título del documento	Informe: Base de datos centralizada		
Nivel de difusión	Público		
Fecha de entrega	30/01/2026		

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Colaboradores del documento

Responsable del documento	CIFAL Málaga		
Redactores del documento	Organización	Revisores	Organización revisora
Martha Goyeneche	CIFAL Málaga	Elena Mesa Alejandro Yankilevich	Fundación CIEDES
Alfiya Urazaeva	CIFAL Málaga		

Descargo de responsabilidad

Financiado por la Unión Europea. Sin embargo, las opiniones y puntos de vista expresados son exclusivamente los del autor o los autores y no reflejan necesariamente los de la Unión Europea ni los de la Agencia Ejecutiva Europea en el Ámbito Educativo y Cultural (EACEA). Ni la Unión Europea ni la autoridad que concede el préstamo se responsabilizan de ellos.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Historial del documento

Formulario	Versión	Fecha	Autor/revisor	Organización
Borrador	1.0	17/01/2026	Martha Goyeneche	CIFAL Málaga
Borrador	2.0	21/01/2026	Alfiya Urazaeva Elena Mesa Alejandro Yankilevich	CIFAL Málaga Fundación CIEDES
Final	3.0	28/01/2026	Martha Goyeneche Alfiya Urazaeva Elena Mesa Alejandro Yankilevich	CIFAL Málaga CIFAL Málaga Fundación CIEDES Fundación CIEDES

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Contenido

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

1. Introducción

1.1. Rol de la base de datos dentro del proyecto ReTo

La base de datos centralizada constituye el núcleo operativo del WP2 y uno de los pilares transversales del proyecto ReTo. Su objetivo es generar evidencia empírica, desagregada y contextualizada sobre discursos y delitos de odio, que sirva como insumo estratégico para el resto de los paquetes de trabajo.

A través de esta herramienta, se busca fortalecer la capacidad analítica del consorcio, facilitar intervenciones focalizadas, y contribuir al diseño de políticas públicas más justas, inclusivas y basadas en datos reales.

Luego de la finalización del proyecto RETO, la herramienta quedará a disposición de las entidades que la requieran, y se podrá ampliar su cobertura a otros territorios.

1.2. Relación con WP2 y otros Work Packages

El desarrollo de la base de datos se enmarca en la Tarea T2.1 del WP2, liderado por la Fundación CIEDES, en coordinación con CIFAL Málaga y Movimiento Contra la Intolerancia.

Este WP actúa como eje vertebrador del proyecto, alimentando con información sistemática y actualizada al WP1 (Coordinación), WP3 (Apoyo a víctimas), WP4 (Periodismo ético), WP5 (Deporte e inclusión) y WP6 (Intervención cultural). Su diseño garantiza la coherencia interna del proyecto y su capacidad de impacto a distintos niveles.

1.3. Alcance temporal: datos 2025–2027

La base de datos centralizada del proyecto ReTo está concebida para abarcar un periodo temporal que se extiende desde el año 2025 hasta junio de 2027, coincidiendo con el horizonte de ejecución del proyecto financiado por la línea CERV de la Unión Europea. Este intervalo ha sido definido estratégicamente para permitir tanto una reconstrucción retrospectiva de tendencias recientes en discursos y delitos de odio, como el seguimiento en tiempo real de su evolución a lo largo del proyecto. De este modo, se combinan análisis de carácter histórico con una capacidad operativa de monitoreo contemporáneo, lo cual amplía considerablemente el potencial analítico y predictivo del sistema.

La incorporación de datos retrospectivos permite establecer líneas base y patrones de comportamiento que servirán como referencia para la identificación de cambios

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

significativos, picos de actividad o correlaciones con eventos sociales, políticos o mediáticos relevantes.

Esta dimensión histórica es clave no solo para dimensionar el fenómeno, sino también para validar la eficacia de las intervenciones implementadas durante el proyecto. Asimismo, el seguimiento en tiempo real facilitará la activación de mecanismos de alerta temprana, capaces de detectar aumentos súbitos en la intensidad o frecuencia de discursos de odio en canales digitales específicos, y permitirá una respuesta más ágil y coordinada entre actores institucionales y comunitarios.

La frecuencia de actualización de la base de datos será diaria, y mensualmente se realizarán los análisis de gráficos y archivo de datos recogidos, lo cual garantiza un equilibrio entre capacidad técnica, calidad del dato y oportunidad analítica. El sistema ha sido diseñado con suficiente flexibilidad como para incorporar actualizaciones más frecuentes en caso de situaciones críticas, eventos extraordinarios o necesidades específicas detectadas por los usuarios del sistema. Esta adaptabilidad será especialmente relevante en momentos de alta sensibilidad social —por ejemplo, durante campañas electorales, crisis migratorias o eventos internacionales—, en los que los discursos de odio suelen intensificarse y diversificarse.

En definitiva, el alcance temporal de la base de datos ha sido definido para maximizar su utilidad analítica y operativa dentro del marco del proyecto ReTo, pero también con una visión de futuro que permita su continuidad y ampliación en caso de consolidarse como herramienta institucional estable para el monitoreo de los discursos y delitos de odio en el entorno digital.

1.4. Enfoque general (interseccional, sensible al género y basado en evidencia)

El diseño metodológico de la base de datos centralizada del proyecto ReTo está fundamentado en un enfoque interseccional, sensible al género y orientado por principios de rigurosidad científica y evidencia empírica. Este enfoque no solo responde a los objetivos estratégicos del proyecto, sino que también se alinea con los compromisos internacionales en materia de derechos humanos, igualdad y lucha contra la discriminación. Al integrar estas perspectivas de forma transversal, se busca construir una herramienta técnicamente robusta, pero también socialmente transformadora, capaz de captar la complejidad estructural y contextual de los discursos y delitos de odio en el espacio digital.

El componente interseccional, permite analizar cómo las distintas formas de opresión —racismo, sexismo, LGTBIQ+fobia, antigitanismo, capacitismo, entre

Cofinanciado por la Unión Europea

otras— no se presentan de forma aislada, sino que se entrecruzan generando efectos múltiples y acumulativos sobre ciertos grupos sociales.

Esta mirada evita las simplificaciones excesivas y permite identificar con mayor precisión los perfiles de riesgo, las dinámicas específicas de exclusión y los mecanismos de reproducción del odio en función de múltiples ejes de desigualdad. En la práctica, esto se traduce en un sistema de categorización flexible y en la incorporación de variables que permiten capturar esta multiplicidad, como el cruce entre identidad étnico-racial, orientación sexual, género, condición migratoria y clase social.

Por su parte, la perspectiva de género se incorpora como principio estructurante en todas las etapas del diseño e implementación del sistema. Esta perspectiva reconoce que las mujeres, así como las personas trans, no binarias y otras identidades del espectro LGTBIQ+, son blanco frecuente de discursos de odio que reproducen estructuras patriarcales y heteronormativas profundamente arraigadas.

Además, el análisis de estos discursos permite identificar no solo ataques explícitos, sino también formas simbólicas de violencia digital que refuerzan estereotipos, silencian voces y perpetúan desigualdades. La base de datos contempla categorías específicas para registrar y analizar estos casos, promoviendo así una mayor visibilidad y capacidad de respuesta ante la violencia de género en el entorno digital.

Finalmente, el proyecto se articula bajo un enfoque basado en evidencia, que garantiza que cada decisión técnica y metodológica esté sustentada en datos empíricos, fuentes verificables y marcos de análisis validados. Este enfoque implica una trazabilidad completa de los procesos de recolección, clasificación y análisis de la información, así como la aplicación de protocolos de validación cruzada entre humanos y modelos de IA.

Se prioriza la transparencia en la construcción de las librerías terminológicas, la documentación técnica abierta y la evaluación continua del desempeño de los algoritmos utilizados. Todo ello enmarcado en un compromiso ético con la protección de los derechos de las personas afectadas, la no revictimización y la protección de datos personales, en conformidad con los estándares del Reglamento General de Protección de Datos (RGPD) y los principios rectores del programa CERV.

En conjunto, este enfoque metodológico no solo eleva la calidad técnica de la base de datos, sino que refuerza su legitimidad política y social como herramienta al servicio de una ciudadanía más informada, empoderada y protegida frente al odio.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

2. Marco Conceptual y Normativo

Este apartado establece los fundamentos conceptuales, jurídicos y metodológicos que sustentan la construcción de la base de datos del proyecto ReTo. Dado que gran parte de este contenido ya ha sido desarrollado en otros entregables, a continuación, se ofrecen introducciones sintéticas y se indica con precisión en qué documentos y páginas se encuentra el desarrollo íntegro de cada subtema.

2.1. Definiciones operativas: discurso de odio y delitos de odio

La base de datos se construye sobre una definición operativa clara y funcional de *discurso de odio* y *delitos de odio*, que permite establecer criterios homogéneos de recolección y clasificación. Se parte de las definiciones proporcionadas por organismos internacionales como la OSCE, ECRI y FRA, adaptadas al contexto andaluz y español.

Consultar en:

- Documento: *RETO_D2.1. ESP (1).pdf*
  Apartado: "2.1. Definiciones clave"
  Página: 5

- Documento: *D3.1. Estrategia de Concienciación Pública...pdf*
  Apartado: "3.1.1. Definición de discurso de odio"
  Página: 16

2.2. Marco normativo: UE, España y Andalucía

El desarrollo de la base de datos se inscribe dentro de un marco jurídico multiescalar. A nivel europeo, se tiene en cuenta la Directiva 2012/29/UE y las recomendaciones de ECRI. En el plano estatal y autonómico, se consideran el Código Penal español, los protocolos de actuación de la Fiscalía y las iniciativas normativas desarrolladas en Andalucía en materia de igualdad y no discriminación.

Consultar en:

- Documento: *RETO_D2.1. ESP (1).pdf*
  Apartado: "2.2. Marco normativo multiescalar"
  Página: 7

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

- Documento: *D3.1. Estrategia de Concienciación Pública...pdf*
  Apartado: "3. Marco jurídico nacional e internacional"
  Páginas: 14–15

2.3. Tipologías de odio integradas en la base de datos

La base de datos recoge un conjunto de tipologías de odio que reflejan tanto las categorías jurídicas reconocidas como las realidades sociales emergentes. Estas incluyen, entre otras: racismo, xenofobia, LGTBIQ+fobia, antigitanismo, islamofobia, antisemitismo, sexismo y odio hacia personas con discapacidad.

Consultar en:

- Documento: *RETO_D2.1. ESP (1).pdf*
  Apartado: "2.3. Tipologías de odio: definición operativa"
  Página: 9

- Documento: *D3.1. Estrategia de Concienciación Pública...pdf*
  Apartado: "2.3. Clasificación de discursos de odio"
  Página: 12

2.4. Referencia a estándares europeos de recogida de datos (OSCE, FRA, ECRI)

La construcción de la base de datos se inspira en estándares internacionales de recolección de datos sobre odio, priorizando la comparabilidad, la desagregación y la inclusión de categorías interseccionales. Se siguen las recomendaciones metodológicas del Observatorio de la OSCE para los Crímenes de Odio, la Agencia de Derechos Fundamentales (FRA) y la Comisión Europea contra el Racismo y la Intolerancia (ECRI).

Consultar en:

- Documento: *RETO_D2.1. ESP (1).pdf*
  Apartado: "2.4. Recomendaciones metodológicas internacionales"
  Página: 10

3. Arquitectura General del Sistema

3.1. Diseño conceptual de la base de datos

El sistema de base de datos desarrollado en el marco del Proyecto ReTo ha sido concebido como una infraestructura integral y flexible para la recolección, almacenamiento, procesamiento y análisis de información relacionada con

Cofinanciado por la Unión Europea
ReTo Red de Tolerancia contra los delitos de odio

discursos y delitos de odio en Andalucía. Su diseño conceptual responde a un enfoque metodológico que combina tres principios clave: la interseccionalidad, el uso de fuentes múltiples y la orientación a evidencia. Esta arquitectura no solo permite captar la complejidad del fenómeno, sino también generar datos útiles para la acción institucional y social.

El sistema está diseñado para integrar contenidos de naturaleza muy diversa —como mensajes publicados en redes sociales, noticias digitales, registros oficiales e informes institucionales— y vincularlos de forma estructurada a categorías analíticas coherentes con los marcos de clasificación europeos, especialmente los definidos por OSCE/ODIHR, FRA y ECRI. Cada entrada en la base de datos incluye información contextual, temporal y geográfica que facilita su análisis dinámico y comparado. Además, se ha previsto un registro completo del ciclo de vida del dato: desde su ingreso hasta su procesamiento, anotación manual y auditoría, lo que garantiza trazabilidad y control de calidad.

En su conjunto, el diseño conceptual de la base de datos aspira a ser más que una estructura de almacenamiento: se concibe como un ecosistema de información crítico para entender los discursos de odio desde una perspectiva compleja, contextualizada y centrada en los derechos humanos.

3.2. Esquema relacional y principales tablas

La base de datos opera sobre una infraestructura PostgreSQL, elegida por su potencia, estabilidad y compatibilidad con sistemas analíticos avanzados. Se ha estructurado en esquemas diferenciados que permiten una organización lógica y segura de los distintos componentes del sistema. El modelo relacional permite representar las múltiples relaciones entre mensajes, categorías, anotaciones, usuarios y fuentes, con un enfoque modular que facilita su ampliación y mantenimiento.

Entre las tablas principales destacan:

- **messages:** contiene los mensajes captados automáticamente desde redes sociales (como X/Twitter, YouTube, etc.) mediante herramientas de scraping. Incluye metadatos asociados como fecha, canal, y ubicación estimada.
- **manual_label_csv / manual_label:** registran las anotaciones realizadas por humanos dentro del proyecto, con indicadores de validez y nivel de acuerdo entre anotadores.
- **labels, label_taxa, label_modifiers:** forman la estructura jerárquica de la taxonomía del discurso de odio, incluyendo categorías primarias (como

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

xenofobia o sexismo), subcategorías y modificadores que permiten ajustar el análisis a niveles de intensidad, tono o ironía.

- **audit_log:** contiene el registro completo de cada anotación, identificando el lote, usuario responsable, fecha y resultado del proceso.

Complementariamente, se integran tablas provenientes de fuentes institucionales como el Ministerio del Interior y la Fiscalía General del Estado, lo que permite realizar análisis comparados entre lo detectado en el entorno digital y los datos oficiales de delitos de odio registrados por el sistema judicial y policial.

3.3. Variables incluidas

El sistema ha sido diseñado para capturar una amplia gama de variables que permitan análisis multicausales y desagregados, en coherencia con el enfoque interseccional del proyecto. Entre las principales variables destacan:

- Metadatos técnicos del mensaje: fecha, hora, canal, idioma, y código de contenido.
- **Target detectado y validado**: grupo o colectivo hacia el que se dirige el discurso, de acuerdo con la taxonomía del proyecto.
- **Tipo de incidente**: diferenciando entre discurso de odio, incitación a la violencia, negacionismo, acoso, etc.
- **Canal** de publicación o diseminación del contenido.
- **Ubicación geográfica estimada**, si está disponible (a partir de metadatos o contexto textual).
- **Intensidad del discurso**: medida mediante un sistema de niveles escalonados, de acuerdo con criterios de gravedad.
- **Uso de humor, ironía o sarcasmo**, como modificadores del sentido del mensaje.
- **Dimensión interseccional**: posibilidad de múltiples targets o categorías cruzadas (por ejemplo, islamofobia con componente de género).
- **Variables institucionales**: categorías penales, tipologías oficiales de delitos de odio, y naturaleza del caso según registros de Interior o Fiscalía.

Esta estructura de variables permite realizar estudios longitudinales, análisis de correlación, visualizaciones segmentadas y monitoreo automatizado, con posibilidad de agregar nuevas variables conforme evolucione el sistema.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

3.4. Protocolos de anonimización y privacidad

Desde el inicio, el sistema ha sido diseñado bajo estrictos principios de protección de datos personales, conforme al Reglamento General de Protección de Datos (RGPD) y a las recomendaciones éticas del programa CERV. Todos los identificadores personales (como nombres de usuarios, IDs de cuentas o direcciones IP) son anonimizados de forma irreversible mediante algoritmos hash SHA-256. Esta técnica garantiza que no es posible reconstruir la identidad del emisor del mensaje, incluso si se accede al contenido o a los metadatos de forma externa.

Además, se aplica un proceso de minimización de datos: solo se almacenan aquellos campos que son estrictamente necesarios para el análisis, descartando cualquier información superflua o potencialmente sensible. Cada carga o modificación de datos queda documentada en el audit_log, incluyendo detalles del lote de procesamiento, herramientas utilizadas, fecha y responsables del tratamiento, lo que permite una trazabilidad completa del flujo de datos y refuerza los mecanismos de auditoría y rendición de cuentas.

Estos protocolos permiten que el sistema pueda ser utilizado con seguridad tanto por investigadores como por entidades públicas sin comprometer la privacidad de las personas cuyos mensajes forman parte del análisis, asegurando así su uso ético, legal y responsable.

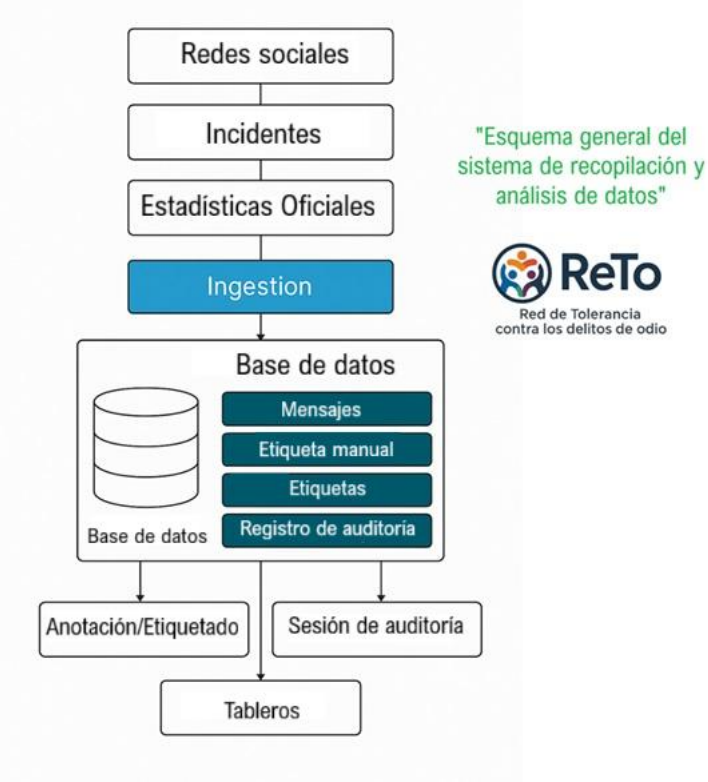


Gráfico 1: "Esquema general del sistema de recopilación y análisis de datos"

Cofinanciado por la Unión Europea

4. Tecnologías Utilizadas

El desarrollo de la base de datos centralizada del Proyecto ReTo se apoya en un ecosistema tecnológico moderno y modular, que combina herramientas de código abierto con soluciones adaptadas a los requerimientos específicos de recolección, procesamiento, análisis y visualización de datos sensibles. La selección tecnológica responde a criterios de eficiencia, sostenibilidad, compatibilidad institucional y facilidad de mantenimiento a largo plazo.

4.1. Lenguajes de programación

El lenguaje de programación principal utilizado es **Python**, por su versatilidad, su ecosistema maduro en tareas de ciencia de datos y procesamiento de lenguaje natural (NLP), así como su amplia comunidad y capacidad de integración con múltiples servicios. Python se emplea en distintas fases del proyecto: para el desarrollo de scripts de **scraping**, en la etapa de **preprocesamiento textual**, en la automatización del pipeline **ETL** (Extract, Transform, Load), y en la orquestación general de procesos de ingestión y etiquetado.

A nivel de gestión y consulta de datos estructurados, se utiliza **SQL** sobre **PostgreSQL**, lo cual permite un control detallado del modelo relacional, consultas complejas optimizadas y compatibilidad con herramientas de visualización y exportación. La estructura del sistema favorece el uso de sentencias SQL avanzadas para análisis estadístico y categorización de patrones del discurso.

4.2. Frameworks para scraping, monitorización y análisis

Para la recolección de contenido desde redes sociales y medios digitales, el proyecto emplea un enfoque híbrido, combinando el uso de APIs oficiales, servicios de scraping gestionado y procesamiento local de datos.

En el caso de YouTube, la obtención de datos se realiza mediante la YouTube Data API, respetando los límites y condiciones de uso de la plataforma. Para la plataforma X, la recolección de mensajes se lleva a cabo a través de batches automatizados generados mediante Apify, utilizando actores especializados que permiten una captura estructurada y escalable del contenido público de interés.

De forma complementaria, para determinadas fuentes web y medios digitales, se emplean bibliotecas como Requests y HTTPX para realizar peticiones HTTP controladas, junto con BeautifulSoup y lxml para el parseo de contenido HTML cuando es necesario.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

En la fase de procesamiento y análisis preliminar, se utilizan Pandas y NumPy para la manipulación estructurada de los datos recolectados. Las expresiones regulares (Regex) se aplican para la identificación de patrones léxicos y el filtrado inicial de mensajes candidatos. Para el análisis lingüístico y la tokenización, se emplea spaCy, permitiendo normalizar los textos y prepararlos para las fases posteriores de etiquetado y análisis.

Esta combinación de herramientas y servicios permite construir un pipeline modular y extensible, capaz de adaptarse a distintas plataformas, cambios tecnológicos y variaciones en el volumen de datos, manteniendo la trazabilidad y el control del proceso de recolección y análisis.

4.3. Infraestructura del servidor y almacenamiento

El sistema opera sobre una **infraestructura Linux** segura, estable y optimizada para procesamiento de datos. El motor de base de datos **PostgreSQL** actúa como núcleo del sistema, con una arquitectura organizada en **esquemas diferenciados** que separan las distintas fases y componentes del flujo de datos (por ejemplo: ingesta, anotación, resultados, auditoría).

El almacenamiento se organiza en **lotes de procesamiento**, lo que permite trazabilidad, control de versiones y depuración de errores. Además, la base de datos ofrece **compatibilidad con exportación en múltiples formatos**, incluyendo **CSV** para interoperabilidad básica y **Parquet** para análisis de alto rendimiento en entornos distribuidos o plataformas de big data. Este diseño permite integraciones futuras con sistemas de terceros sin pérdida de granularidad ni consistencia estructural.

4.4. API externas utilizadas

Para enriquecer el análisis y ampliar la capacidad de extracción de datos desde plataformas digitales, se integran varias **APIs externas** con funciones específicas:

YouTube Data API (Google): permite acceder a información estructurada sobre vídeos, comentarios, canales y metadatos, facilitando la recolección automatizada de contenido relevante sobre discursos de odio en esta plataforma.

Google Perspective API (cuando se aplica): proporciona un análisis preliminar de la "toxicidad" de los comentarios o mensajes mediante modelos de aprendizaje automático entrenados en corpus multilingües. Aunque su uso es exploratorio y requiere validación local, ofrece una referencia externa para evaluar el impacto emocional o violento del discurso.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Estas APIs son utilizadas de forma controlada, conforme a los términos de servicio de cada proveedor y los protocolos éticos del proyecto.

4.5. Sistemas de visualización

En el componente de visualización, se han integrado herramientas adaptadas a los distintos perfiles de usuario del sistema. Por un lado, **Power BI** se utiliza para generar **informes institucionales de carácter formal**, con dashboards estructurados que permiten analizar tendencias, filtrar por tipologías de odio, territorios o periodos de tiempo, y generar reportes exportables en formatos estándar para toma de decisiones.

Por otro lado, **Looker Studio (antes Google Data Studio)** se emplea como plataforma de visualización **ágil e interactiva**, especialmente útil para exploración interna del sistema, validación participativa de categorías y análisis rápido del rendimiento de las etiquetas manuales y automáticas. Ambas herramientas están conectadas mediante APIs o archivos intermedios que extraen datos directamente desde la base PostgreSQL o desde exportaciones controladas, manteniendo la integridad del dato original.

Este enfoque dual garantiza tanto el rigor necesario para procesos institucionales como la flexibilidad para el trabajo técnico y social colaborativo.

Tecnología / Herramienta	Descripción	Uso principal en ReTo	Licencia
Python	Lenguaje de programación interpretado, ampliamente usado en ciencia de datos y automatización	Scraping, procesamiento de texto, automatización del pipeline ETL, análisis estadístico	Open Source (PSF License)
SQL / PostgreSQL	Sistema de gestión de bases de datos relacional de alto rendimiento	Almacenamiento estructurado, consultas, modelado relacional, exportaciones	Open Source (PostgreSQL License)

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Tecnología / Herramienta	Descripción	Uso principal en ReTo	Licencia
Requests / HTTPX	Librerías de Python para realizar solicitudes HTTP	Recolección de datos desde plataformas web (scraping)	Open Source (Apache 2.0 / BSD)
BeautifulSoup / lxml	Librerías para parseo y limpieza de HTML/XML	Extracción de contenido de páginas web	Open Source (MIT / BSD)
Pandas / NumPy	Librerías para manejo eficiente de datos estructurados y análisis numérico	Manipulación de datos, normalización, cálculo de estadísticas	Open Source (BSD)
Regex	Expresiones regulares de Python	Filtrado de patrones léxicos y preprocesamiento textual	N/A (lenguaje base)
spaCy	Biblioteca de procesamiento de lenguaje natural (NLP)	Tokenización, segmentación, anotación semántica inicial	Open Source (MIT)
YouTube Data API	API oficial de Google para extracción de datos desde YouTube	Recolección de vídeos, comentarios, canales y metadatos	Gratuita con registro (uso limitado bajo Términos de Google)
Google Perspective API	API de análisis de contenido desarrollada por Jigsaw (Google)	Evaluación preliminar de "toxicidad" y tono en comentarios textuales	Gratuita con límites (Licencia de uso Google AI)
Power BI	Plataforma de análisis empresarial de Microsoft	Visualización estructurada de datos, informes institucionales	Propietaria (licencia institucional o

Cofinanciado por la Unión Europea

Tecnología / Herramienta	Descripción	Uso principal en ReTo	Licencia
			gratuita limitada)
Looker Studio	Herramienta de visualización de datos online (antes Google Data Studio)	Exploración ágil de datos y visualización participativa	Gratuita (Google Services License)
Linux (Ubuntu/Debian)	Sistema operativo de código abierto, estable y seguro	Infraestructura del servidor, entorno de ejecución	Open Source (varias licencias GNU/Linux)
CSV / Parquet	Formatos de archivo para exportación e interoperabilidad	Exportación de datos, integración con otros sistemas	Open (estándar industrial)

Tabla 1: Resumen de tecnologías utilizadas

5. Procesos de Recolección de Datos

La fase de recolección de datos constituye el punto de partida del sistema centralizado del Proyecto ReTo, y ha sido diseñada para garantizar rigurosidad metodológica, diversidad de fuentes, legalidad en la obtención de los datos y relevancia analítica respecto a los discursos de odio en el contexto andaluz. La estrategia contempla tanto fuentes digitales abiertas como registros oficiales institucionales, aplicando filtros y protocolos de calidad que aseguren la trazabilidad, pertinencia y seguridad de los datos tratados. Los datos se recolectarán y analizarán durante el tiempo que dure el proyecto. Actualmente estamos en la fase inicial de construcción, puesta en marcha, validación manual y aprendizaje de la IA para la sistematización total.

5.1. Fuentes online (redes y medios digitales)

La recolección de datos online se centra exclusivamente en contenidos públicos generados en plataformas digitales a partir de publicaciones realizadas por medios de comunicación previamente seleccionados. En el caso de redes sociales, se

Cofinanciado por la Unión Europea

analizan los comentarios asociados a dichos contenidos, sin realizar capturas de perfiles personales ni conversaciones privadas.

Para YouTube, se utilizan los comentarios públicos asociados a vídeos publicados por los medios seleccionados, obtenidos mediante la API oficial de la plataforma. En el caso de X, los mensajes se recolectan a partir de publicaciones de medios y sus interacciones públicas, utilizando procesos de scraping gestionado mediante Apify.

Este enfoque no pretende representar a la población general ni medir el nivel real de discurso de odio en la sociedad, sino constituir una ventana de observación sobre la manifestación de discursos hostiles en entornos mediáticos digitales. Los resultados obtenidos no son extrapolables a la totalidad de la ciudadanía, sino que deben interpretarse como indicadores exploratorios dentro de un contexto comunicativo específico.

5.2. Fuentes institucionales

Como complemento a la recolección digital, el sistema integra datos provenientes de **fuentes oficiales institucionales**, principalmente:

- El **Ministerio del Interior**, a través de sus informes anuales sobre incidentes relacionados con delitos de odio en España, los cuales desglosan estadísticas por motivación, territorio, grupo afectado y evolución temporal.
- La **Fiscalía General del Estado**, mediante sus memorias anuales, que incluyen referencias a casos judicializados, líneas de acción del Ministerio Fiscal y evolución en la aplicación de normativas penales en materia de odio y discriminación.

Estos datos son procesados y adaptados al modelo de la base, permitiendo **análisis comparativos** entre lo registrado en plataformas digitales y lo judicializado o reconocido oficialmente. Esta doble fuente permite identificar posibles discrepancias entre el odio denunciado y el odio detectado en línea, ofreciendo insumos clave para el diagnóstico institucional y la incidencia política.

5.3. Criterios de calidad y filtrado

Tras la recolección inicial, los datos pasan por un proceso de normalización y control de calidad. Se eliminan duplicados, se verifica el idioma de los mensajes y se descartan aquellos contenidos que no cumplen con los criterios mínimos de análisis.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

En esta fase se aplican filtros léxicos basados en la librería de términos del proyecto y en reglas lingüísticas diseñadas para reducir ruido y falsos positivos. Este filtrado no clasifica automáticamente los mensajes como discurso de odio, sino que permite priorizar aquellos contenidos con mayor probabilidad de relevancia para su revisión manual posterior.

5.4. Frecuencia de actualización

La actualización de los datos se realiza con una periodicidad diaria, ajustada a las limitaciones técnicas y a las políticas de uso de las plataformas digitales y APIs empleadas. En el caso de X, la generación de nuevos batches mediante Apify se planifica según el volumen de actividad detectado en los medios monitorizados. Para YouTube, la frecuencia de actualización se ajusta a las cuotas disponibles de la API oficial y al ritmo de publicación de los canales analizados.

Este esquema flexible permite mantener un flujo de datos constante, garantizando la continuidad del análisis sin comprometer la estabilidad técnica del sistema.

5.5. Limitaciones y riesgos identificados

El proceso de recolección enfrenta una serie de **limitaciones técnicas y riesgos operativos** que han sido identificados y abordados mediante estrategias de mitigación específicas. Entre los principales riesgos destacan:

- **Restricciones en las APIs** utilizadas, tanto en términos de volumen permitido como en los cambios imprevisibles en sus políticas de acceso.
- **Cambios estructurales en las plataformas digitales** (por ejemplo, rediseños de interfaces o bloqueos de automatización), que pueden afectar temporalmente la capacidad de scraping.
- **Variaciones significativas en el volumen y tono de los comentarios recolectados**, directamente relacionadas con la coyuntura informativa, lo que genera picos de actividad difíciles de anticipar.

Para mitigar estos riesgos, el sistema incorpora:

- Protocolos de **anonimización estricta y automatizada**, que garantizan la legalidad del tratamiento de datos.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

- Mecanismos de **auditoría técnica y humana** de los lotes recolectados, con verificación cruzada y registro de errores.
- **Documentación completa del pipeline de recolección**, desde la fuente hasta el procesamiento final, que permite replicabilidad, depuración y transparencia.

Estas estrategias aseguran que el sistema sea resiliente, ético y capaz de sostenerse operativamente durante todo el ciclo del proyecto, incluso en contextos cambiantes.

Gráfico 2: "Proceso de recolección de datos".

6. Sistema de Clasificación y Etiquetado

6.1. Metodología de etiquetado (basada en Hatemedia y CRISP-DM)

El proceso de etiquetado del proyecto ReTo se apoya en dos marcos metodológicos: el manual Hatemedia como referencia principal y el modelo CRISP-DM, un estándar internacional para proyectos de minería de datos que estructura el proceso en seis

Cofinanciado por la Unión Europea

fases (comprensión del problema, comprensión de los datos, preparación, modelado, evaluación y despliegue). En este proyecto, CRISP-DM orienta la organización coherente de la recopilación, transformación y etiquetado.

La metodología utilizada incluye:

1. Preprocesamiento del texto: normalización, tokenización, limpieza lingüística, detección inicial de coincidencias según el diccionario Hatemedia normalizado y el diccionario complementario de términos generales de hostilidad del proyecto ReTo, e identificación preliminar del posible target.

2. Pre-etiquetado automático: aplicación de la librería combinada (Hatemedia + diccionario ReTo), detección preliminar de términos relevantes y filtrado de mensajes candidatos.

3. Etiquetado humano experto: los anotadores clasifican odio / no odio / dudoso, asignan categoría (una de las seis definidas), así como intensidad y humor cuando corresponde.

4. Consolidación y trazabilidad: integración en PostgreSQL y registro mediante audit_log para reconstruir todo el flujo de cada mensaje (origen, reglas aplicadas, anotaciones realizadas y revisiones posteriores).

6.2. Librería de términos: ampliaciones y validaciones

La librería utilizada parte del diccionario Hatemedia (7.122 términos), normalizado y optimizado según análisis de uso real en medios andaluces, junto con una lista de stopwords específicas para evitar falsos positivos.

Sobre esta base se ha incorporado un diccionario complementario de términos generales de hostilidad desarrollado en el proyecto ReTo (insultos directos, metáforas deshumanizantes, hostilidad ideológica y contra profesiones o roles públicos, etc.), que permite capturar formas de discurso hostil no cubiertas por Hatemedia.

En fases posteriores está previsto documentar versiones formales de la librería (Hatemedia + ReTo) e incorporar un sistema de versionado que permita reproducir con precisión las condiciones de cada fase de etiquetado.

6.3. Niveles de intensidad del discurso

El proyecto adopta una escala simplificada de 1 a 3, aplicable solo cuando el mensaje ha sido clasificado como ODIO:

1 – Leve: ironía, desdén o formas de hostilidad moderada.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

2 – Ofensivo: insultos claros o estigmatización explícita.

3 – Hostilidad/incitación: deshumanización, deseo de daño o llamadas a exclusión o expulsión.

Para los mensajes clasificados como NO ODIO o DUDOSO no se asigna intensidad. La intensidad final se determina exclusivamente mediante etiquetado humano.

6.4. Categorías de odio y clasificación del target

El sistema de clasificación del proyecto ReTo se organiza en seis categorías principales de discurso de odio. Estas categorías se implementan en formato normalizado (snake_case) para facilitar su uso en el pipeline de datos:

1. Odio_etnico_cultural_religioso: Se refiere a expresiones de hostilidad, estigmatización o incitación al odio dirigidas contra personas o grupos en función de su **origen étnico, nacionalidad, religión, cultura o prácticas culturales asociadas**. Esta categoría incluye discursos racistas, xenófobos, islamófobos, antisemitas, anticristianos en contextos minoritarios, y ataques a símbolos culturales colectivos.

Ejemplos: “Que se vayan a su país”, “Los musulmanes no respetan nuestras leyes”, “Todos los latinos son delincuentes”.

2. Odio genero_identidad_orientacion: Agrupa mensajes que expresan odio, desprecio o violencia simbólica hacia personas por razón de su **género, identidad de género u orientación sexual**, incluyendo además ataques al feminismo, a los derechos LGTBIQ+ o a colectivos trans. Esta categoría cubre misoginia, homofobia, lesbofobia, transfobia y discursos antiigualdad.

Ejemplos: “Las feminazis arruinan todo”, “Eso te pasa por ser maricón”, “Las trans nunca serán mujeres”.

3. Odio_condicion_social_economica_salud: Incluye discursos discriminatorios basados en la **pobreza, la exclusión social, el nivel educativo, la discapacidad, enfermedades mentales o físicas, y situaciones de dependencia**. Este tipo de odio suele manifestarse mediante lenguaje deshumanizante, criminalización de la pobreza o burla hacia condiciones de salud.

Ejemplos: “Todos los vagabundos son drogadictos”, “Con lo que cuesta mantener a los discapacitados”, “Ese loco debería estar encerrado”.

4. Odio_ideologico_politico: Recoge expresiones de odio motivadas por la **adscripción política, creencias ideológicas, participación en movimientos sociales o percepciones sobre afiliación partidaria**. Aunque esta categoría es

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

más difusa y requiere interpretación contextual, permite analizar la polarización política y la hostilidad contra determinadas corrientes ideológicas.

Ejemplos: "Todos los comunistas deberían desaparecer", "La ultraderecha es una enfermedad social", "Los votantes de X son basura".

5. Odio_personal_generacional: Esta categoría abarca discursos basados en **la edad (edadismo)** —ya sea hacia personas mayores o jóvenes— y aquellos que atacan características personales no ligadas directamente a condiciones médicas o discapacidades reconocidas. También incluye burlas o desprecios por apariencia física.

Ejemplos: "Los viejos no deberían opinar", "Estos niñatos no saben de nada", "Qué asco de gorda".

6. Odio_profesiones_roles_publicos *(categoría auxiliar):* Esta categoría recoge ataques dirigidos específicamente contra personas que ejercen **profesiones públicas o roles de responsabilidad institucional**, como periodistas, personal sanitario, docentes, jueces, policías o representantes políticos. Aunque en algunos casos estos ataques pueden tener componentes ideológicos, su inclusión diferenciada responde a la necesidad de monitorear **formas de odio hacia lo público y el sistema institucional**.

Ejemplos: "Los médicos son unos asesinos por vacunar", "Los periodistas mienten por dinero", "Los políticos deberían morir todos".

6.5. Validación cruzada humano–IA

El proyecto no desarrolla aún un modelo automático completo, sino un sistema de pre-etiquetado automatizado que actúa como filtro inicial. Su función es reducir el volumen de mensajes a revisar, identificar candidatos relevantes y facilitar el trabajo de los anotadores. La validación consiste en comparar el pre-etiquetado con la etiqueta humana, medir precisión, revisar falsos positivos y ajustar criterios.

6.6. Control de sesgos y calidad del etiquetado

El control de calidad se estructura en tres dimensiones:

(1) Medidas procedimentales:

- Uso de Hatemedia como base metodológica, complementado con el diccionario ReTo de términos generales de hostilidad.

- Homogeneización de criterios entre anotadores mediante el Manual de Etiquetado actualizado.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

(2) Medidas técnicas:

- Anonimización completa mediante hashing.

- Registro de cada lote de procesamiento mediante audit_log, permitiendo identificar cuándo ingresó un mensaje, con qué reglas fue procesado y quién realizó la anotación.

(3) Medidas analíticas:

- En esta primera fase se aplican controles básicos: revisión de coherencia interna, detección manual de inconsistencias y análisis simple de dispersión de intensidad. Métricas avanzadas, como el acuerdo interanotador, se incorporarán más adelante.

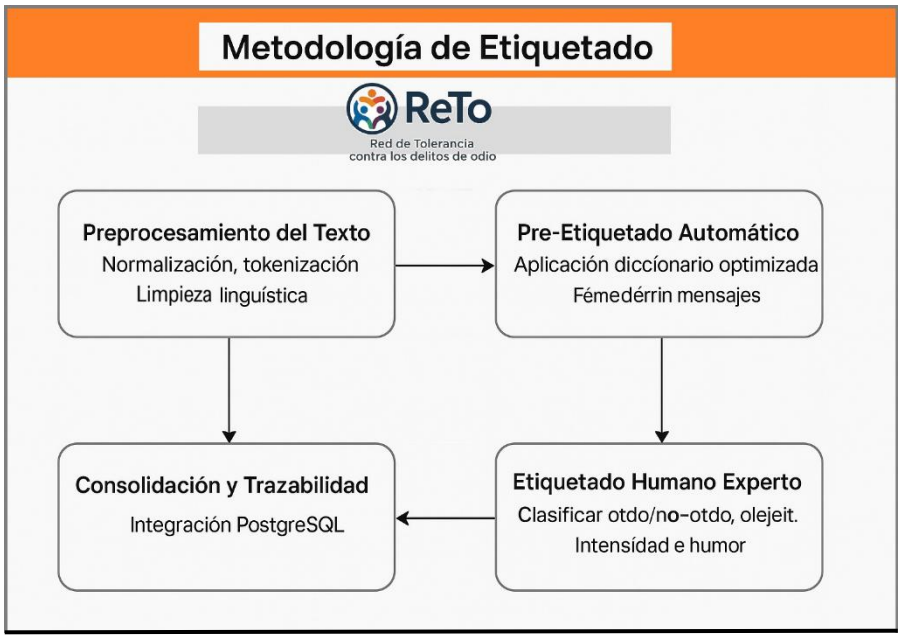


Gráfico 3: Metodología de etiquetado.

6.7. Bases de datos en proceso de etiquetado y validación (enlaces trabajo actual)

El proyecto dispone actualmente de dos bases de datos activas en proceso de análisis y validación, alojadas de forma temporal en hojas de cálculo compartidas en Google Drive, utilizadas como entorno operativo para el trabajo colaborativo de etiquetado y control de calidad.

Base de datos YouTube (etiquetado manual completo)

La base de datos correspondiente a YouTube contiene mensajes recolectados a partir de comentarios públicos asociados a vídeos publicados por medios de comunicación previamente seleccionados.

Cofinanciado por la Unión Europea

Esta base ha sido etiquetada íntegramente por anotadores humanos, siguiendo los criterios definidos en el Manual de Etiquetado del proyecto, e incluye la identificación de discurso de odio, su categorización y variables asociadas cuando corresponde.

Actualmente, esta base constituye el conjunto de referencia para la validación metodológica del sistema y para la evaluación de herramientas de apoyo automatizado.

El acceso a esta base se realiza mediante el siguiente enlace:

https://docs.google.com/spreadsheets/d/1XRkeNE9rq0O7feE3az7TnX-Wbmv-fWn15Cjl4AVA0To/edit?usp=sharing

Base de datos X (validación humana de pre-etiquetado automático)

La base de datos correspondiente a la plataforma X contiene mensajes recolectados a partir de publicaciones e interacciones públicas vinculadas a medios de comunicación definidos previamente. En este caso, los mensajes han sido pre-etiquetados mediante un sistema de apoyo basado en modelos IA de lenguaje (LLM), con el objetivo de priorizar contenidos potencialmente relevantes. Dicho pre-etiquetado es posteriormente validado y corregido por anotadores humanos, garantizando que la decisión final sobre la clasificación del contenido sea siempre de carácter humano.

Esta base se encuentra actualmente en fase de validación manual y permite evaluar la precisión del pre-etiquetado automático, así como ajustar criterios y umbrales de priorización.

El acceso a esta base se realiza mediante el siguiente enlace:

https://docs.google.com/spreadsheets/d/1mTavDecRfdWsNqHQT8Pe3kghF5B9xrR36ISCO9GVwno/edit?usp=sharing

Infraestructura y migración prevista

Ambas bases de datos se encuentran alojadas de forma temporal en Google Drive como solución operativa para el trabajo colaborativo durante las fases de etiquetado y validación.

Está previsto que, en un plazo aproximado de 40 días, con fecha de implementación estimada para el 27 de febrero, estas bases ya se encuentren migradas a una infraestructura centralizada en un servidor PostgreSQL, lo que permitirá mejorar la trazabilidad, el versionado, la escalabilidad del sistema y su integración con los procesos analíticos y de explotación posteriores.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

7. Inteligencia Artificial y Métricas de Evaluación

El uso de técnicas de inteligencia artificial (IA) en el Proyecto ReTo responde a un enfoque prudente, progresivo y enfocado en la mejora de eficiencia operativa sin sustituir el juicio humano en la toma de decisiones analíticas. Durante esta primera fase del proyecto, la IA se emplea exclusivamente para asistir en el **pre-etiquetado automatizado** del contenido, actuando como un filtro inicial que permite **reducir el volumen de mensajes** a ser procesados manualmente. Esta aproximación garantiza control sobre el sistema, minimiza sesgos y facilita la validación cruzada entre herramientas automáticas y experticia humana.

7.1. Modelos aplicados

En esta etapa inicial, el sistema de IA se basa en técnicas de procesamiento automático del lenguaje natural (PLN) y análisis léxico de baja complejidad. No se han implementado aún modelos basados en aprendizaje profundo (deep learning) ni procesos de representación semántica por vectorización (como embeddings), dado que el foco está en establecer un sistema trazable, replicable y alineado con los recursos disponibles.

(a) Modelos no supervisados

Se utilizan principalmente modelos **no supervisados** para realizar tareas de limpieza, agrupación exploratoria y detección basada en reglas. Entre las técnicas aplicadas destacan:

- **Matching léxico avanzado**: uso de expresiones regulares y patrones contextuales para identificar la presencia de términos clave o estructuras lingüísticas asociadas a discursos hostiles.
- **Preprocesamiento del texto**: incluye tokenización, lematización, eliminación de duplicados y normalización lingüística para mejorar la precisión de los filtros automáticos.
- **Clustering exploratorio**: se han realizado pruebas limitadas de agrupación de mensajes para identificar núcleos temáticos o correlaciones frecuentes, sin que estos influyan directamente en la clasificación final. Esta técnica se considera preliminar y de apoyo para futuras fases analíticas.

(b) Modelos supervisados

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Actualmente, **no se han implementado modelos supervisados** entrenados con datos etiquetados. El sistema no realiza clasificación automatizada con algoritmos como árboles de decisión, redes neuronales, SVM o similares. Todo el proceso de etiquetado definitivo —incluyendo la categoría, el target y el nivel de intensidad— es realizado exclusivamente por anotadores humanos expertos, con apoyo de guías metodológicas y revisión continua.

7.2. Entrenamiento y datasets utilizados

La automatización parcial del sistema se basa en tres insumos fundamentales:

1. **Dataset de comentarios de YouTube (ReTo)**
    - Contiene mensajes recolectados exclusivamente de **perfiles oficiales de medios andaluces** en YouTube.
    - Cada mensaje es sometido a procesos de **normalización lingüística, anonimización irreversible** y filtrado por idioma y relevancia.
    - Este dataset es dinámico, ampliándose con cada ciclo semanal de recolección y ajustado según la evolución del vocabulario hostil detectado.
2. **Diccionario Hatemedia optimizado**
    - Se parte de la versión original del diccionario Hatemedia, con **7.122 términos**, y se realiza una curaduría orientada al uso local.
    - Se eliminan entradas no pertinentes, se corrigen ambigüedades y se amplía la base con expresiones hostiles emergentes detectadas durante la recolección.
    - Aunque la versión actual está en fase experimental, se prevé implementar un sistema de versionado formal que documente los cambios léxicos por fecha y fuente.
3. **Dataset de etiquetas manuales**
    - Representado por la tabla manual_label_csv e integrado en PostgreSQL, este archivo contiene los resultados del proceso de etiquetado humano con las siguientes variables:
        - **Categoría** (tipo de discurso de odio)
        - **Target** (grupo afectado)

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

- **Intensidad** (escala 0–3)
- **Uso de humor/ironía**

- Este conjunto se utilizará como base para entrenamientos futuros de modelos supervisados, una vez se alcance el volumen necesario y se validen las condiciones éticas y técnicas para su uso.

En esta fase, **no se han generado variables derivadas** más allá de la normalización textual, aunque se contempla incorporar vectores semánticos, análisis de polaridad y representaciones embeddings en futuras iteraciones.

7.3. Métricas de rendimiento

La evaluación del sistema se realiza mediante un conjunto de métricas que permiten valorar tanto la **eficacia del pre-etiquetado automático** como la **coherencia y consistencia del etiquetado humano**. Estas métricas no se entienden como indicadores absolutos, sino como instrumentos de mejora continua del sistema.

1. Métricas del pre-etiquetado automático

- **Precisión**: mide qué porcentaje de los mensajes marcados como "candidatos relevantes" por el sistema automático contenían efectivamente elementos hostiles o problemáticos según el juicio humano.
- **Tasa de reducción de volumen**: indica cuánto se logra reducir el corpus a revisar gracias al prefiltrado automatizado, sin comprometer la detección de casos significativos.
- **Tasa de falsos positivos y negativos**: se realiza una revisión manual sobre una muestra aleatoria de mensajes pre-etiquetados para estimar errores del sistema, lo que permite ajustar los filtros y el diccionario.

2. Métricas del etiquetado humano

- **Consistencia interna**: se verifica que las etiquetas aplicadas mantengan coherencia lógica entre categoría, target e intensidad, y que no existan combinaciones contradictorias.
- **Distribución de intensidad (escala 0–3)**: se analiza cómo se distribuyen los niveles asignados, detectando posibles sesgos de sobre-etiquetado (por exceso de gravedad) o sub-etiquetado (por exceso de prudencia).
- **Coherencia interanotador**: aunque aún no se aplican métricas cuantitativas formales como el **Kappa de Cohen**, se realiza una comparación

Cofinanciado por la Unión Europea
ReTo Red de Tolerancia contra los delitos de odio

cruzada sobre pequeños lotes para identificar divergencias sistemáticas. Este proceso ha permitido refinar la guía de anotación y mejorar el entrenamiento de los equipos humanos.

- Este enfoque de evaluación integral refuerza el compromiso del proyecto con una ciencia de datos socialmente responsable, técnicamente válida y éticamente sólida.

Gráfico 4: Inteligencia artificial y métricas de evaluación.

8. Arquitectura del Panel de Visualización

El sistema de visualización de datos del Proyecto ReTo ha sido concebido como un componente estratégico para la comunicación de hallazgos, la toma de decisiones institucionales y la incidencia pública. Su diseño responde a los principios de accesibilidad, escalabilidad y utilidad práctica, permitiendo a diversos perfiles de

Cofinanciado por la Unión Europea

usuarios —instituciones, periodistas, OSC, académicos— explorar de forma ágil los datos recopilados, etiquetados y analizados durante el proyecto.

El panel puede desplegarse en dos entornos complementarios: **Looker Studio**, por su capacidad de exploración dinámica y gratuita, y **Power BI**, por su solidez en la generación de informes estructurados y su integración con entornos corporativos. Ambos se alimentan desde una **base de datos PostgreSQL**, a través de **vistas específicas** que filtran, transforman y agregan los datos relevantes.

La arquitectura general del panel se estructura en cuatro capas principales:

1. Capa de datos (ingreso y organización del contenido)

En esta primera capa se integran las diferentes fuentes de información procesadas por el sistema:

- **Mensajes anonimizados** provenientes de redes sociales, especialmente YouTube, extraídos desde perfiles de medios andaluces.
- **Etiquetas manuales** aplicadas por el equipo de anotadores, que incluyen categoría de odio, target, intensidad e indicadores de humor o ambigüedad.
- **Resultados del filtrado automático**, como marcadores de pre-etiquetado, intensidad estimada y presencia de términos clave.
- **Tablas oficiales** obtenidas del Ministerio del Interior y de la Fiscalía General del Estado, estandarizadas para su comparación con los datos digitales.
- **Taxonomía del discurso de odio y metadatos asociados**, como la fecha, medio, ubicación aproximada, y tipo de canal donde fue emitido el mensaje.

Esta capa constituye la **base estructural** sobre la cual se realizan las operaciones posteriores de análisis y visualización.

2. Capa de transformación (procesamiento analítico intermedio)

A partir de la base de datos central, se generan **vistas especializadas** en PostgreSQL que permiten transformar los datos brutos en indicadores analíticos procesables. Algunas de las vistas clave incluyen:

- **v_message_anonymized**: versión depurada de los mensajes con metadatos mínimos, apta para consulta pública.
- **v_labels_extended:** contiene etiquetas cruzadas con información contextual, como medio, fecha y geolocalización aproximada.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

- **v_stats_by_year:** ofrece agregaciones por año, categoría, intensidad y target.

Estas vistas permiten generar **agregaciones dinámicas** en función de múltiples variables: tipo de odio, intensidad del discurso, grupo afectado, plataforma de origen, y período temporal. Esta capa constituye el **puente entre los datos estructurados y la interfaz visual**, garantizando rendimiento, seguridad y flexibilidad.

3. Capa de visualización (interfaz de usuario y análisis interactivo)

En esta capa se despliega la información mediante una interfaz gráfica interactiva y personalizable, con las siguientes características:

- **Filtros dinámicos** que permiten seleccionar el año, el medio de comunicación, la categoría de odio, el target o la intensidad del discurso.
- **Indicadores clave de análisis**, como el porcentaje de mensajes hostiles sobre el total, la distribución por niveles de intensidad (0 a 3), y la frecuencia relativa de targets específicos.
- **Series temporales** que permiten observar la evolución del discurso de odio por categoría, medio o intensidad.
- **Gráficos comparativos**, tanto en formato de barras como circulares o líneas, que facilitan la interpretación rápida de tendencias y correlaciones.

La visualización se adapta a distintos dispositivos y puede configurarse por perfiles de usuario, ofreciendo una experiencia accesible para actores técnicos e institucionales.

4. Capa de salida (exportación y comunicación)

El sistema permite la **exportación de informes y visualizaciones** en diversos formatos (PDF, Excel, imagen), adaptados a las necesidades de:

- **Prensa**: datos segmentados y visualmente comprensibles para acompañar coberturas sobre odio digital.
- **Autoridades públicas**: informes técnicos con agregados por territorio y grupo vulnerable, útiles para políticas públicas.
- **Organizaciones sociales y comunitarias (OSC)**: acceso a datos estructurados para monitoreo ciudadano, incidencia y sensibilización.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Esta capa refuerza el compromiso del proyecto con la **apertura, la transparencia y la utilidad pública**, facilitando la circulación de evidencia en formatos comprensibles, relevantes y basados en datos.

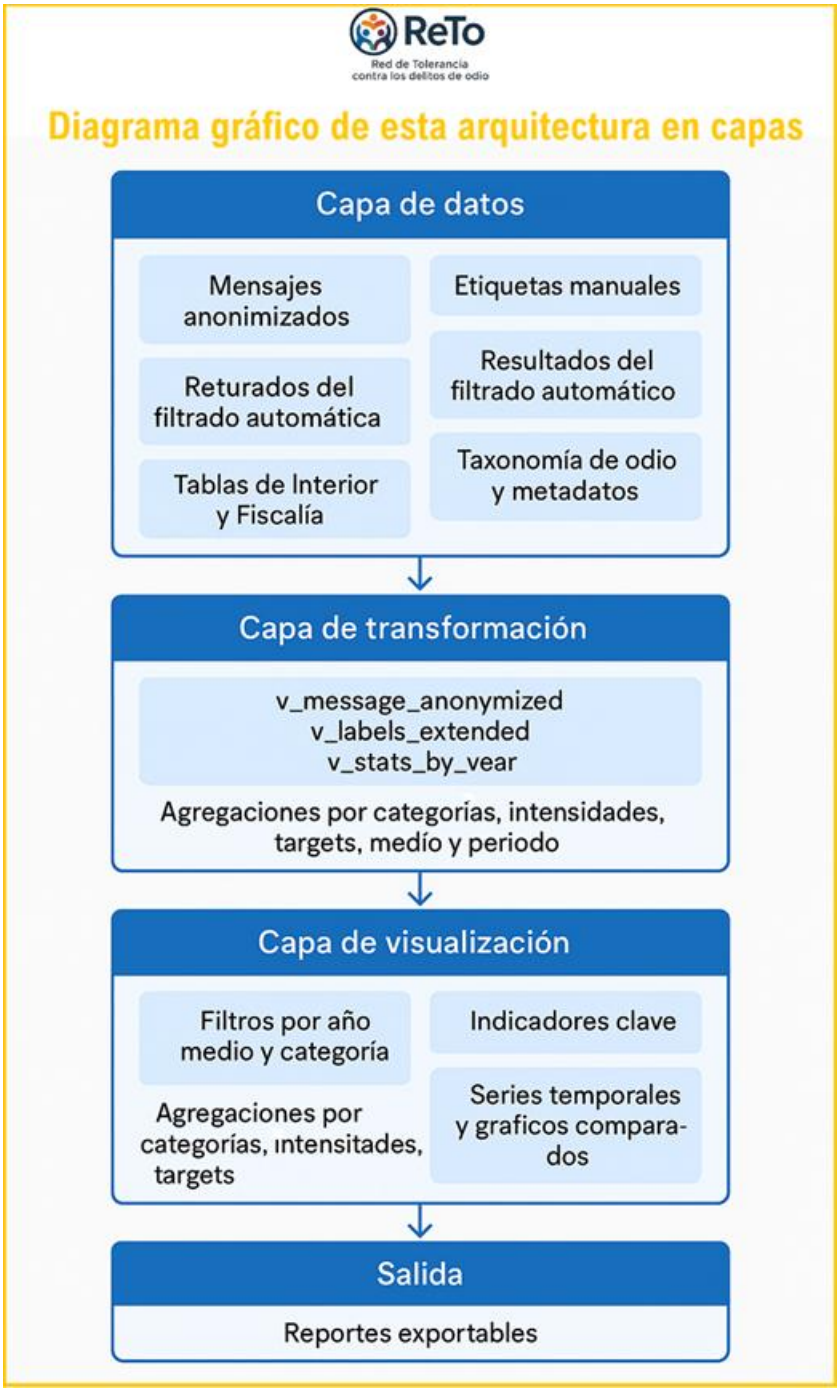


Gráfico 5: Diagrama Gráfico de arquitectura en capas.

Cofinanciado por la Unión Europea

9. Validación Participativa

El Proyecto ReTo reconoce que la construcción de una base de datos sobre discursos y delitos de odio no puede limitarse a un enfoque exclusivamente técnico o académico. Por el contrario, requiere incorporar de forma estructurada y sistemática las voces de quienes enfrentan, analizan o intervienen directamente en contextos de odio: **organizaciones sociales, periodistas, operadores institucionales y cuerpos policiales**. Por ello, se ha desarrollado una estrategia de **validación participativa** que acompaña el proceso técnico en todas sus fases, garantizando su pertinencia contextual, su sensibilidad cultural y su utilidad práctica.

9.1. Actores implicados (OSC, periodistas, cuerpos policiales)

En esta primera etapa, se ha involucrado a un conjunto diverso de actores clave con presencia y experiencia en el territorio andaluz:

- **Organizaciones de la sociedad civil (OSC)** que trabajan con población migrante, personas LGTBIQ+, comunidades gitanas, víctimas de violencia de género y otros colectivos expuestos a discursos de odio. Estas entidades han sido fundamentales para validar la taxonomía, aportar ejemplos reales de mensajes, y señalar nuevas categorías emergentes.
- **Periodistas y medios de comunicación**, especialmente aquellos vinculados al Colegio de Periodistas de Andalucía, han contribuido a definir qué medios y perfiles son más representativos del entorno digital andaluz, y han aportado perspectivas críticas sobre la relación entre discurso mediático y polarización social.
- **Representantes de cuerpos policiales y operadores jurídicos**, como policías locales, miembros de unidades especializadas y fiscalías de delitos de odio, han ofrecido retroalimentación sobre las categorías legales, las formas en que el odio se judicializa y las limitaciones del sistema actual de denuncias.

Esta diversidad de actores ha permitido un enfoque **multidimensional**, combinando saberes técnicos, comunitarios, jurídicos y comunicacionales.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

9.2. Resultados de talleres o consultas previas

Durante los primeros meses del proyecto, se llevaron a cabo **espacios de consulta semiestructurada y talleres participativos**, organizados en colaboración con los socios del consorcio ReTo. Algunas de las conclusiones extraídas incluyen:

- Necesidad de **visibilizar discursos ambiguos o "normalizados"** que, sin ser explícitamente hostiles, refuerzan estigmas o exclusiones estructurales.
- Importancia de **refinar las categorías interseccionales**, especialmente en casos en que un mismo mensaje afecta simultáneamente a mujeres migrantes, personas trans o minorías religiosas.
- Relevancia de incorporar el **humor como categoría analítica**, ya que muchos discursos de odio se difunden bajo la apariencia de ironía o sátira.
- Observaciones críticas sobre el uso de determinadas etiquetas, que podrían generar revictimización o malinterpretaciones si no se aplican con precisión y contexto.

Estos hallazgos no solo informaron el ajuste de los criterios de etiquetado, sino que fortalecieron el compromiso del proyecto con una ciencia de datos socialmente situada y políticamente consciente.

9.3. Ajustes realizados a partir de la retroalimentación

Como resultado directo de estas consultas, se implementaron una serie de **ajustes concretos en la arquitectura y el funcionamiento del sistema**, entre ellos:

- Rediseño parcial de la **guía de anotación**, incorporando ejemplos aportados por las OSC y validando expresiones propias del contexto andaluz.
- Revisión del vocabulario base del diccionario Hatemedia optimizado, eliminando términos con doble sentido o escasa relevancia local.
- Inclusión de un campo específico para marcar **uso de humor, sarcasmo o ironía**, lo que permite un análisis más matizado.
- Creación de un módulo de comentarios internos dentro del sistema de anotación, para registrar dudas, decisiones argumentadas y casos límite.
- Acondicionamiento de la interfaz de visualización (dashboard) con filtros ajustados a las necesidades de análisis de las OSC y entidades colaboradoras.

Estos cambios mejoran la **legitimidad del sistema**, su aplicabilidad en contextos reales y su alineación con los objetivos del programa CERV.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

9.4. Próximas fases de revisión

La validación participativa se concibe como un **proceso continuo**, no como un evento puntual. En los próximos meses se prevé:

- La **organización de nuevos talleres de retroalimentación**, centrados en la evaluación del etiquetado humano y en la usabilidad del panel de visualización.
- La **consulta específica con comunidades afectadas**, asegurando que sus voces se integren de forma directa y no solo representadas por terceros.
- El **pilotaje de indicadores de impacto**, evaluando en qué medida el sistema mejora la capacidad de acción de instituciones públicas, OSC y ciudadanía frente al odio.
- La **documentación del proceso participativo** como parte de los entregables del proyecto, contribuyendo a la transparencia y a la replicabilidad del modelo.

Este enfoque asegura que la base de datos no solo sea técnicamente robusta, sino también políticamente legítima y socialmente útil.

10. Escalabilidad, Mantenimiento y Futuro

10.1. Roadmap hasta junio de 2027

El desarrollo de la base de datos se ha planificado para avanzar en paralelo con el resto de entregables del proyecto ReTo. El roadmap prevé:

- **2025 (completo):** diseño conceptual, definición metodológica y construcción del esquema relacional.
- **Primer semestre 2026:** integración de fuentes digitales e institucionales, implementación de scraping automatizado y pruebas con modelos NLP supervisados.
- **Segundo semestre 2026:** validación participativa con actores clave (OSC, medios, policía), optimización del sistema de visualización y ajustes metodológicos.
- **2027:** expansión territorial, formación de usuarios institucionales y comunitarios, y producción del dashboard final interoperable con otros sistemas nacionales y europeos.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Este cronograma está alineado con los entregables y metas definidas en el *Grant Agreement*, particularmente en relación con los objetivos del WP2 y el impacto previsto en el WP6 (*Sustainability and Transferability*).

10.2. Replicabilidad en otras regiones de España

El diseño modular de la base de datos —basado en una arquitectura relacional flexible, un sistema de clasificación validado y una librería terminológica adaptable— permite su implementación en otros territorios con diferentes características sociopolíticas. Esta replicabilidad está prevista como una fase estratégica posterior a su consolidación en Andalucía, permitiendo, además, su integración con políticas públicas autonómicas y nacionales.

10.3. Sostenibilidad técnica tras el proyecto

Con el objetivo de asegurar su sostenibilidad más allá de la duración del proyecto, se han previsto varias estrategias:

- Licenciamiento abierto de la arquitectura general del sistema y del vocabulario controlado.
- Transferencia de capacidades técnicas a actores institucionales a través de formaciones específicas.
- Alojamiento en infraestructuras compatibles con servidores públicos o académicos.
- Mantenimiento mínimo asegurado mediante documentación técnica y buenas prácticas de desarrollo.

Estas estrategias están en consonancia con las directrices del Programa CERV sobre sostenibilidad de resultados y acceso abierto.

10.4. Recomendaciones para el WP final

Con miras a garantizar la sostenibilidad técnica, política y social de la base de datos centralizada más allá de la duración del proyecto ReTo, se considera imprescindible que se incorpore una serie de medidas estructurales que aseguren su mantenimiento, ampliación y transferencia. En primer lugar, se recomienda el desarrollo de un **plan de transferencia tecnológica** que contemple la cesión progresiva del sistema a entidades públicas o académicas que puedan hacerse cargo de su operatividad una vez finalizado el marco de financiación CERV. Este plan

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

deberá incluir sesiones de formación técnica para perfiles diversos —desde técnicos informáticos hasta personal analista y responsables institucionales—, así como la producción de materiales de capacitación específicos y adaptados a los diferentes niveles de uso.

11. Resumen Ejecutivo

11.1. Objetivo del informe

El presente informe tiene como finalidad documentar de forma exhaustiva el proceso de diseño, implementación y validación de la base de datos centralizada desarrollada en el marco del Proyecto ReTo, específicamente dentro del *Work Package 2* (WP2), titulado "Investigación y análisis de datos sobre delitos de odio". Esta base de datos constituye uno de los productos técnicos y estratégicos más relevantes del proyecto, dado que proporciona la infraestructura necesaria para la recolección sistemática, el análisis estructurado y la visualización interactiva de información empírica relacionada con discursos y delitos de odio en el entorno digital, con un enfoque aplicado inicialmente al contexto andaluz, pero diseñado desde su origen con capacidad de escalabilidad nacional e incluso europea.

El informe detalla los fundamentos conceptuales, metodológicos y técnicos que sustentan el desarrollo de esta herramienta, así como los avances alcanzados en su implementación hasta la fecha. Se describen las variables incorporadas, la arquitectura del sistema, los modelos de análisis utilizados, los protocolos de recolección de datos y los mecanismos de validación técnica y participativa. A través de esta descripción, se busca no solo dar cuenta del estado actual del sistema, sino también proporcionar insumos útiles para su replicación, evaluación y mejora continua en fases posteriores del proyecto o en otras iniciativas similares.

El diseño de la base de datos está guiado por un enfoque interseccional, sensible al género y basado en evidencia, lo que implica una atención especial a las múltiples formas de discriminación que afectan a personas y colectivos en situación de vulnerabilidad, especialmente en contextos de violencia digital. Este enfoque permite no solo capturar la complejidad del fenómeno, sino también generar información útil para la toma de decisiones políticas, el diseño de estrategias preventivas y la implementación de intervenciones más efectivas y justas. En ese sentido, la base de datos no se limita a ser una herramienta de diagnóstico, sino que se posiciona como una infraestructura crítica para la acción pública y social, orientada a fortalecer la respuesta institucional y comunitaria frente al aumento de los discursos de odio en Europa.

Este informe, por tanto, se dirige tanto a los socios del consorcio como a actores públicos, técnicos y comunitarios interesados en comprender cómo se construye y

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

utiliza una herramienta de este tipo, cuáles son sus fundamentos éticos y técnicos, y qué posibilidades abre para el desarrollo de políticas basadas en datos y centradas en los derechos humanos.

11.2. Estado actual del desarrollo de la base de datos

A enero de 2026, la base de datos se encuentra en fase avanzada de desarrollo. Ya se han completado las etapas de diseño conceptual, estructuración del esquema relacional y definición de variables clave, incorporando dimensiones como el tipo de incidente, el grupo diana, el canal de difusión, la intensidad del discurso, la interseccionalidad y la geolocalización. Se espera tener una sistematización y entrenamiento de las herramientas de IA para la búsqueda de datos que no necesite de un etiquetado manual, sino que trabaje de forma autónoma y con una fiabilidad del 80% en el primer trimestre de 2026.

Se ha iniciado la integración de datos provenientes de fuentes digitales públicas y sociales, que incluyen tanto plataformas en línea (como son redes sociales y medios de comunicación digitales), como portales institucionales oficiales (como los del Ministerio del Interior o la Fiscalía General del Estado). Para su tratamiento se aplican tecnología de scraping y modelos de procesamiento del lenguaje natural.

El estudio sobre discurso de odio se centrará en un listado de medios de comunicación de la Comunidad Autónoma de Andalucía, incluyendo sus 8 provincias, y las redes sociales que manejan los mismos medios de comunicación. Es decir, que no se evaluarán las redes sociales a nivel general, sino las cuentas vinculadas a los respectivos medios de comunicación.

En cuanto a la recolección de datos basada en los delitos de odio, se trabajará a nivel nacional, ya que básicamente se utilizarán fuentes oficiales del Ministerio del Interior del Gobierno de España y de la Fiscalía General.

11.3. Principales avances y próximos pasos

El desarrollo de la base de datos centralizada del Proyecto ReTo ha experimentado avances sustanciales en múltiples frentes, consolidando una infraestructura técnica y metodológica que sienta las bases para su uso efectivo como herramienta de análisis, monitoreo e intervención. Uno de los hitos más relevantes ha sido la validación del enfoque metodológico basado en el modelo CRISP-DM (Cross-Industry Standard Process for Data Mining), que ha permitido estructurar el proceso de construcción de la base de datos de manera ordenada, iterativa y orientada a la mejora continua. Este modelo ha guiado desde la comprensión del problema hasta

Cofinanciado por la Unión Europea

la preparación y modelado de datos, facilitando la integración entre los objetivos analíticos del proyecto y sus necesidades operativas.

Otro avance clave ha sido la adaptación y ampliación de la librería terminológica desarrollada originalmente en el proyecto europeo Hatemedia, la cual ha sido contextualizada a la realidad andaluza mediante procesos de curaduría lingüística, análisis discursivo y consulta con expertos. Esta librería constituye el corazón semántico del sistema de clasificación de discursos de odio, y ha sido complementada con términos emergentes, expresiones locales y categorías sensibles a las múltiples formas de discriminación, incluyendo antigitanismo, islamofobia, LGTBIQ+fobia, misoginia, aporofobia, edadismo y discursos contra personas migrantes, entre otros.

Asimismo, se ha completado una primera versión funcional de los dashboards interactivos, los cuales permiten visualizar la información recopilada a través de gráficos dinámicos, mapas georreferenciados, líneas de tiempo y filtros por categoría, canal o intensidad del discurso. Estas herramientas de visualización no solo facilitan la interpretación de los datos por parte de equipos técnicos, sino que también habilitan su uso en procesos de incidencia política, sensibilización pública y formación institucional.

De cara a los próximos meses, se proyectan una serie de acciones estratégicas para consolidar y expandir el sistema. En primer lugar, se dará prioridad a la automatización de los procesos de recolección y actualización de datos, mediante la integración de técnicas avanzadas de scraping, procesamiento del lenguaje natural (NLP) y aprendizaje automático. Esto permitirá aumentar la eficiencia del sistema, reducir errores manuales y mejorar la capacidad de respuesta ante fenómenos emergentes.

En segundo lugar, se avanzará en una fase intensiva de validación participativa, que implicará la colaboración activa con organizaciones de la sociedad civil, periodistas, cuerpos policiales y otros actores institucionales. Esta validación se centrará en evaluar la pertinencia de las categorías, la claridad de las visualizaciones, la utilidad del sistema para distintos perfiles de usuario, y la identificación de posibles sesgos en los modelos de clasificación automática.

Finalmente, esta herramienta quedará preparada para realizar una escalabilidad territorial, orientada a la implementación del sistema en otras comunidades autónomas españolas.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

12. Conclusiones

- El diseño e implementación de la base de datos centralizada del Proyecto ReTo representa una contribución estratégica tanto para el análisis riguroso de los discursos y delitos de odio en Andalucía como para la mejora de los mecanismos de prevención, respuesta institucional y protección de los derechos de los colectivos más vulnerables.

  Esta iniciativa se inserta en un contexto social, político y tecnológico complejo, donde las manifestaciones de odio no solo se multiplican en el espacio digital, sino que adoptan formas cada vez más sofisticadas, fragmentadas y normalizadas. Frente a este escenario, el desarrollo de herramientas tecnológicas con base empírica, capaces de identificar, clasificar y visualizar estos fenómenos en tiempo real, se vuelve no solo necesario, sino urgente.

- A lo largo del proceso, el equipo técnico y metodológico ha consolidado una arquitectura de datos robusta, flexible y orientada a la acción, que recoge tanto la dimensión cuantitativa como cualitativa del fenómeno. La base de datos integra variables clave como la categoría del discurso, el grupo objetivo, el canal de difusión, la intensidad, la interseccionalidad y la georreferenciación, lo que permite no solo registrar incidentes, sino también contextualizarlos y analizarlos en su complejidad.

  Esta riqueza estructural es el resultado de un enfoque metodológico que combina estándares internacionales (como los definidos por la OSCE, ECRI o FRA), marcos normativos estatales y autonómicos, y aportes epistemológicos provenientes del feminismo interseccional, los estudios críticos de raza y los derechos humanos.

- Uno de los principales logros ha sido la adaptación del sistema a las particularidades del territorio andaluz, no solo desde el punto de vista técnico, sino también sociocultural.

  La identificación de términos y expresiones propias del contexto, la validación de casos reales anonimizados, y la construcción de una librería dinámica sensible a nuevas formas de odio digital, han permitido calibrar el sistema para que responda a las realidades locales sin perder su capacidad de articulación nacional y europea. En ese sentido, la base de datos no solo es replicable, sino también escalable, lo cual abre un horizonte de trabajo colaborativo con otras comunidades autónomas e incluso con organismos multilaterales.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

- Otro aspecto fundamental es la incorporación progresiva de modelos de inteligencia artificial supervisados y no supervisados para el procesamiento del lenguaje natural, lo cual permite automatizar tareas de clasificación y detección de patrones discursivos.

  Si bien estos modelos aún se encuentran en fase de prueba y optimización, su desempeño preliminar es prometedor y abre la puerta a procesos de análisis predictivo, alertas tempranas y monitoreo automatizado con bajo margen de error. Estos avances técnicos, sin embargo, han sido acompañados de una reflexión crítica sobre los límites de la automatización y la necesidad de mecanismos de control humano que aseguren interpretaciones contextualizadas, éticas y culturalmente pertinentes.

- Ahora bien, el proceso no ha estado exento de desafíos. Uno de los principales ha sido la disponibilidad y acceso a fuentes institucionales de calidad, especialmente en lo que respecta a datos actualizados y desagregados provenientes del sistema judicial, las fuerzas de seguridad o la fiscalía.

  A ello se suma la dificultad para integrar fuentes con diferentes formatos, niveles de estructuración o protocolos de anonimización. En el ámbito técnico, el desafío ha sido equilibrar precisión analítica con escalabilidad, sin perder de vista la protección de datos y los marcos regulatorios europeos (como el RGPD).

- En este contexto, se desprenden varias recomendaciones clave para las siguientes etapas del proyecto. En primer lugar, es necesario seguir fortaleciendo alianzas con instituciones públicas, tanto para garantizar el flujo de datos como para consolidar canales de uso práctico de la base de datos en procesos de denuncia, prevención y reparación.

- En segundo lugar, debe profundizarse el trabajo con la sociedad civil organizada, no solo como fuente de validación empírica, sino también como agente activo en la interpretación y difusión de los datos, en una lógica de empoderamiento ciudadano y rendición de cuentas institucional.

- En tercer lugar, será fundamental consolidar un equipo técnico y metodológico estable que acompañe la evolución del sistema en el tiempo, garantizando su sostenibilidad más allá del horizonte de 2026. Esta sostenibilidad no debe pensarse solo en términos de mantenimiento técnico, sino también como una apuesta ética, política y pedagógica por construir sistemas de información al servicio de la justicia social.

- En suma, la base de datos centralizada no debe concebirse únicamente como un producto técnico del proyecto ReTo, sino como una plataforma viva,

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

transformadora y comprometida, capaz de contribuir activamente a la construcción de sociedades más justas, inclusivas y libres de odio. Su continuidad y expansión dependerán del compromiso colectivo del

13. Lista de gráficos y tablas

14. Glosario (tecnicismos)

API (Application Programming Interface)
Interfaz que permite que diferentes sistemas o aplicaciones informáticas se comuniquen entre sí de forma automatizada. En el proyecto ReTo se utilizan APIs externas (como YouTube Data API o Google Perspective API) para extraer datos estructurados de plataformas digitales.

Anonimización
Proceso técnico mediante el cual se eliminan o transforman los datos personales de manera irreversible para impedir la identificación de las personas. En la base de datos ReTo se aplica mediante algoritmos hash (SHA-256) conforme al RGPD.

Audit log
Registro técnico que documenta todas las acciones realizadas sobre los datos (ingreso, modificación, etiquetado, revisión). Permite garantizar trazabilidad, control de calidad y rendición de cuentas.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Base de datos relacional
Modelo de base de datos que organiza la información en tablas relacionadas entre sí mediante claves. PostgreSQL es el sistema relacional utilizado en el proyecto.

Big data
Conjunto de tecnologías y métodos orientados al tratamiento de grandes volúmenes de datos, caracterizados por su variedad, velocidad y volumen. El diseño del sistema permite futuras integraciones con entornos de big data.

Clustering
Técnica de análisis de datos que agrupa elementos similares sin etiquetas previas. En el proyecto se usa de forma exploratoria para detectar patrones temáticos.

CRISP-DM (Cross-Industry Standard Process for Data Mining)
Estándar metodológico internacional para proyectos de minería de datos que estructura el trabajo en seis fases: comprensión del problema, comprensión de los datos, preparación, modelado, evaluación y despliegue.

Dashboard
Panel visual interactivo que permite explorar y analizar datos mediante gráficos, filtros y tablas dinámicas. En ReTo se implementa mediante Power BI y Looker Studio.

Dataset
Conjunto estructurado de datos utilizado para análisis o entrenamiento de modelos. En el proyecto se manejan datasets de comentarios digitales, etiquetas manuales y fuentes institucionales.

Deep learning
Subcampo del aprendizaje automático basado en redes neuronales profundas. En esta fase del proyecto no se aplica, priorizando modelos más simples y trazables.

Embeddings
Representaciones vectoriales del lenguaje que capturan relaciones semánticas entre palabras o textos. Su incorporación se contempla para fases futuras del proyecto.

ETL (Extract, Transform, Load)
Proceso técnico que consiste en extraer datos de distintas fuentes, transformarlos (limpieza, normalización) y cargarlos en una base de datos centralizada.

Framework
Conjunto de herramientas, bibliotecas y buenas prácticas que facilitan el desarrollo de software. En el proyecto se emplean frameworks y librerías de Python para scraping, análisis y NLP.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Hash (SHA-256)
Función criptográfica que transforma datos en una cadena irrepetible de longitud fija. Se utiliza para anonimizar identificadores personales de forma irreversible.

Interseccionalidad
Enfoque analítico que reconoce la superposición de múltiples formas de discriminación (género, origen, religión, orientación sexual, clase social, etc.) en una misma persona o colectivo.

Label / Etiqueta
Clasificación asignada a un mensaje para identificar si contiene discurso de odio, su categoría, target e intensidad.

Looker Studio
Herramienta de visualización de datos de Google (antes Data Studio) utilizada para exploración interactiva y validación participativa de resultados.

Machine learning (Aprendizaje automático)
Conjunto de técnicas que permiten a un sistema aprender patrones a partir de datos. En ReTo se emplea de forma limitada y asistida, sin sustituir el juicio humano.

Matching léxico
Técnica que identifica coincidencias entre textos y listas de términos predefinidos (diccionarios), utilizada para el pre-etiquetado automático.

Metadatos
Información contextual sobre un dato (fecha, canal, idioma, ubicación aproximada, etc.) que facilita su análisis y clasificación.

NLP / PLN (Natural Language Processing / Procesamiento del Lenguaje Natural)
Conjunto de técnicas informáticas destinadas a analizar y procesar lenguaje humano. Se utiliza para tokenización, normalización y análisis preliminar de textos.

Pipeline
Cadena automatizada de procesos que transforma los datos desde su recolección hasta su análisis y visualización.

PostgreSQL
Sistema de gestión de bases de datos relacional de código abierto, robusto y escalable, utilizado como núcleo del sistema ReTo.

Pre-etiquetado automático
Fase inicial en la que el sistema identifica mensajes potencialmente problemáticos antes de su revisión humana.

Cofinanciado por la Unión Europea

ReTo
Red de Tolerancia contra los delitos de odio

Regex (Expresiones regulares)
Lenguaje de patrones utilizado para buscar, filtrar o validar textos. Se emplea para detectar estructuras lingüísticas asociadas al odio.

Scraping
Técnica automatizada de extracción de información desde páginas web o plataformas digitales, respetando límites legales y técnicos.

Stopwords
Palabras muy frecuentes (como artículos o preposiciones) que se excluyen del análisis textual para mejorar la precisión.

Target
Grupo o colectivo al que va dirigido un discurso (por ejemplo, personas migrantes, mujeres, comunidad LGTBIQ+).

Tokenización
Proceso que divide un texto en unidades mínimas (palabras o tokens) para su análisis computacional.

Visualización de datos
Representación gráfica de información mediante gráficos, mapas y tablas para facilitar su interpretación y uso en la toma de decisiones.

15. Bibliografía

Colegio Profesional de Periodistas de Andalucía (CPPA). Registro Oficial de Medios Digitales de Andalucía ROMDA. https://periodistasandalucia.es/registro-medios-digitales-andalucia-romda/

Hatemedia, Observatorio Digital de Odio. https://hatemedia.es/

Collins, P. H., & Bilge, S. (2016). *Intersectionality*. Polity Press.

Comisión Europea contra el Racismo y la Intolerancia. (2015). *General policy recommendation No. 15 on combating hate speech*. Consejo de Europa.

Consejo de Europa. (2016). *Recommendation CM/Rec(2016)15 on combating hate speech*. Consejo de Europa.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex. *University of Chicago Legal Forum*, 1989(1), 139–167.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

European Commission. (2019). *Ethics guidelines for trustworthy AI*. High-Level Expert Group on Artificial Intelligence.

European Union Agency for Fundamental Rights. (2012). *Making hate crime visible in the European Union: Acknowledging victims' rights*. FRA.

European Union Agency for Fundamental Rights. (2023). *Hate crime recording and data collection practice across the EU*. FRA.

Few, S. (2012). *Show me the numbers: Designing tables and graphs to enlighten* (2nd ed.). Analytics Press.

-Fiscalía General del Estado. (2023). *Memoria anual de la Fiscalía General del Estado*. Gobierno de España.

-Ministerio del Interior. (2023). *Informe sobre incidentes relacionados con los delitos de odio en España*. Gobierno de España.

-OSCE Office for Democratic Institutions and Human Rights. (2014). *Hate crime data collection and monitoring: A practical guide*. OSCE.

-United Nations. (2019). *United Nations strategy and plan of action on hate speech*. Naciones Unidas.

16. Anexos

A. Librería de términos actualizada (ejemplo)
B. Ejemplos anonimizados de entradas reales (Ejemplo)
C. Diagrama de arquitectura y flujo de datos (ejemplo)
D. Listado de medios de comunicación de Andalucía incluidos.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

ANEXO 1 – Extracto de la Librería de Términos del Proyecto ReTo

Estructura de la librería (campos principales)

En su versión actual, cada entrada de la librería incluye, como mínimo, los siguientes campos:

term: forma textual tal y como puede aparecer en el mensaje.
lemma: forma canónica o raíz sobre la que se agrupan variantes.
category: foco principal de hostilidad según la taxonomía ReTo

(p. ej. odio_etnico_cultural_religioso, odio_ideologico_politico, etc.).

subtype (opcional): matiz semántico (insulto genérico, animalización, estereotipo, etc.).
source: procedencia (https://hatemedia.es/Hatemedia, ampliacion_reto, normalizacion_reto, etc.).
notes (opcional): aclaraciones para interpretar el uso del término.

Ejemplo ilustrativo de estructura y términos incluidos en la librería utilizada para el pre-etiquetado del proyecto.

term	lemma	category	subtype	source	notes
moro	moro	odio_etnico_cultural_religioso	estereotipo	Hatemedia	Requiere contexto.

Cofinanciado por la Unión Europea

moros de mierda	moro	odio_etnico_cultural_religioso	insulto_compuesto	ampliacion_reto	Secuencia frecuente.
mena	mena	odio_etnico_cultural_religioso	etiqueta_colectivo	ampliacion_reto	Estigmatización de menores.
sudaca	sudaca	odio_etnico_cultural_religioso	insulto_etnico	Hatemedia	Peyorativo hacia latinoamericanos.
gitano de mierda	gitano	odio_etnico_cultural_religioso	insulto_compuesto	ampliacion_reto	Hostilidad explícita.
maricón	maricón	odio_genero_identidad_orientacion	insulto_homofobo	Hatemedia	Si va dirigido → odio.
bollera	bollera	odio_genero_identidad_orientacion	insulto_homofobo	Hatemedia	
travelo	travesti	odio_genero_identidad_orientacion	insulto_transfobo	ampliacion_reto	
lisiado	lisiado	odio_condicion_social_economica_salud	estereotipo_discap	Hatemedia	
retrasado	retrasado	odio_condicion_social_economica_salud	insulto_capacitista	Hatemedia	
parásitos	parásito	odio_condicion_social_economica_salud	animalizacion	ampliacion_reto	

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

vagos subsidiados	vago	odio_condicion_social_economica_salud	estereotipo_social	ampliacion_reto	
rojos de mierda	rojo	odio_ideologico_politico	insulto_politico	ampliacion_reto	
facha	facha	odio_ideologico_politico	insulto_politico	Hatemedia	
progres de mierda	progre	odio_ideologico_politico	insulto_politico	ampliacion_reto	
vieja decrépita	viejo	odio_personal_generacional	insulto_edadista	ampliacion_reto	
niñato	niñato	odio_personal_generacional	estereotipo_edad	Hatemedia	
inútil de bata	inútil	odio_profesiones_roles_publicos	insulto_profesion	ampliacion_reto	
prensa manipuladora	prensa	odio_profesiones_roles_publicos	deslegitimacion	ampliacion_reto	
periodistas vendidos	periodista	odio_profesiones_roles_publicos	deslegitimacion	ampliacion_reto	

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

ANEXO 2 – Anonimización de Datos

Este anexo muestra, mediante datos completamente simulados, cómo funciona el proceso de anonimización irreversible aplicado en el proyecto ReTo para proteger cualquier dato identificable.

1. Registro antes de la anonimización (simulado)

Campo	Valor original (simulado)
user_id	@usuario123
user_name	María González
message_id	18592033488
fecha_publicación	2025-02-16 13:44:22
mensaje	Estos menas son unos inútiles peligrosos.
url_post	https://twitter.com/...

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

2. Resultado después de la anonimización

Campo	Valor anonimizado
user_hash	3f1a9c0b8cb2de408de8c7e2fbb09e23df7fbc71d47343f03fb7c3be9ec4cdaa
message_hash	b8f3026fbf4b4e7f9dd93dc16bf1499a9b5936c34477f07e52d8e31b77d1445f
fecha_publicación	2025-02-16 13:44:22
mensaje	Estos menas son unos inútiles peligrosos.
metadata	canal = X, medio = simulado, lote = 2025_02_16_01

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

3. Esquema del proceso de anonimización

Identificadores reales
↓ (eliminación)
Campos sensibles eliminados
↓ (hashing irreversible)
Identifiers → SHA-256
↓
Validación de estructura
↓
Inserción del registro anonimizado en PostgreSQL
↓
Registro de auditoría (audit_log)

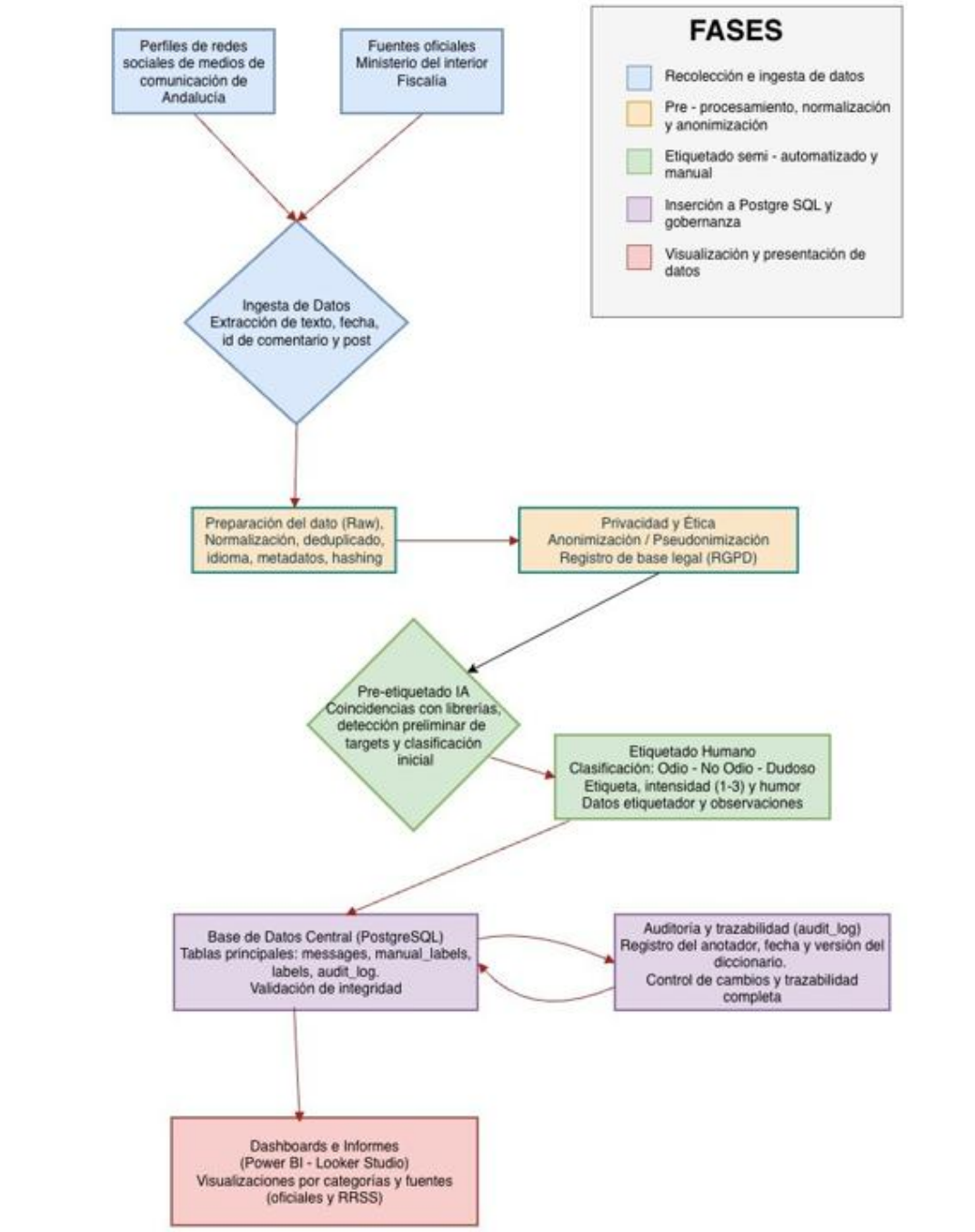
4. Garantías del proceso

- El hashing aplicado es irreversible.
- Nunca se almacena ningún identificador original.
- Solo se conservan texto y metadatos necesarios.
- El audit log documenta el proceso sin conservar identidades.
- El sistema sigue principios de privacy by design.

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

ANEXO 3 Flujo de datos del proyecto ReTo



Cofinanciado por la Unión Europea
ReTo Red de Tolerancia contra los delitos de odio

Explicación del flujo de datos

Las etapas iniciales (1–2) permiten reducir el volumen y filtrar mensajes candidatos. Las etapas centrales (3–4) aseguran calidad y trazabilidad mediante el etiquetado humano y el uso del módulo *audit_log*. Finalmente, los datos se integran en dashboards para análisis y reportes periódicos (5).

¿Qué es audit_log y por qué es esencial en el sistema?

El módulo *audit_log* es un componente crítico de trazabilidad dentro del sistema ReTo. Registra todas las acciones relevantes del ciclo de vida de un mensaje y su proceso de etiquetado. Esto garantiza transparencia metodológica, reproducibilidad y control de calidad.

Cada registro en *audit_log* incluye:

• Hash del mensaje anonimizado
• Identificador del anotador
• Fecha y hora exacta de etiquetado
• Decisión de clasificación (Odio / No Odio / Dudoso)
• Categoría asignada y nivel de intensidad
• Versión de la librería de términos utilizada
• Resultado del pre-etiquetado y observaciones

Su función principal es asegurar que todo el proceso pueda auditarse sin almacenar ningún dato personal, cumpliendo así los principios de privacidad y transparencia.

Cofinanciado por la Unión Europea

ANEXO 4 Medios digitales en proceso de análisis

Listado completo de medios de comunicación andaluces incluidos en el WP2 del proyecto ReTo.

Total de medios a analizar y sus redes en Youtube y X (233 Andalucía)

Medio	Página web
101 TV Antequera	101tv.es
101 TV Marbella	101tv.es
101 TV Málaga	101tv.es
101 TV Ronda	101tv.es
101 TV Sevilla	101tv.es
101 TV Vélez Málaga	101tv.es
101 TV ádiz	101tv.es
101TV (Málaga, Sevilla, Cádiz, Antequera)	101tv.es
7 TV Marbella	7tvandalucia.es
7 TV Málaga	7tvandalucia.es
7TV Andalucía (Red de emisoras)	7tvandalucia.es
8 Directo	8directo.com
8Directo (Algeciras, La Línea, etc.)	8directo.com
9 La Loma	9laloma.tv
ABC Andalucía	abc.es
ABC Andalucía (versión digital regional)	abc.es
ABC de Sevilla	sevilla.abc.es
ANoticias	agenciadenoticias.es
AZ Costa del Sol	azcostadelsol.com
Agencia EFE	efe.com
Ahora Granada	ahoragranada.com
Ahora Mairena	ahoramairena.es
Aionsur	aionsur.com

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Al Sol de la Costa	alsoldelacosta.com
Algeciras al Minuto	algecirasalminuto.com
Alhaurín de la Torre	alhaurindelatorre.com
Aljarafe Digital	aljarafedigital.com
Almería Hoy	almeriahoy.com
Almería Noticias	almerianoticias.es
Almería Press	almeripress.es
Almuñecar Digital	almunecardigital.com
Andalucia Información - Alaclá la Real	andaluciainformacion.es
Andalucía Información	andaluciainformacion.es
Atlas	atlas-news.com
Axarquía TV	axarquiatelevision.es
AxarquíaPlus	axarquiaplus.es
Betis TV	realbetisbalompie.es
COPE Andalucía	cope.es
Cadena SER Andalucía	cadenaser.com
Campiña Digital	campinadigital.me
Canal Coín Radio	canalcoin.com
Canal Luz	canalluztelevision.es
Canal Málaga RTV	canalmalaga.es
Canal San Roque	tv.multimediasanroque.com
Canal Sur Noticias	canalsur.es
Canal Sur Radio / Fiesta / RAI	canalsur.es
Canal Sur TV / Andalucía TV	canalsur.es
Canalcosta TV	canalcostatv.es
Candil Radio (Huércal de Almería)	candilradio.com
Colpisa	colpisa.com
Condavisión	condavision.es
Cordópolis	cordopolis.eldiario.es
Costa Digital	costadigital.es
Costa Noroeste TV (Sanlúcar)	costanoroestetv.com

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Cádiz Buenas Noticias	cadizbuenasnoticias.com
Cádiz Directo	cadizdirecto.com
Córdoba Buenas Noticias	cordobabn.com
Córdoba Hoy	cordobahoy.es
D-Cerca	d-cerca.com
Diario Avanza	diarioavanza.es
Diario Axarquía	diarioaxarquia.com
Diario Bahia de Cádiz	diariobahiadecadiz.com
Diario Bahía deCádiz	diariobahiadecadiz.com
Diario Costa	diariocosta.com
Diario Córdoba	diariocordoba.com
Diario Guadalquivir	diarioguadalquivir.com
Diario Jaén	diariojaen.es
Diario Sexitano	diariosexitano.com
Diario Sur	diariosur.es
Diario de Almería	diariodealmeria.es
Diario de Cádiz	diariodecadiz.es
Diario de Jerez	diariodejerez.es
Diario de Morón	diariodemoron.com
Diario de Sevilla	diariodesevilla.es
Diario de la Línea	diariodelalinea.es
Diario Área	diarioarea.com
Doce TV (Vivacable)	visovision.com
Dos Hermanas Diario Digital	doshermanasdiariodigital.com
El Castillo de San Fernando	elcastillodesanfernando.es
El Confidencial Andaluz	elconfidencialandaluz.com
El Correo de Andalucía	elcorreoweb.es
El Día de Córdoba	eldiadecordoba.es
El Esañol de Sevilla	elespanol.com
El Español de Málaga	elespanol.com
El Estrecho Digital	elestrechodigital.com

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

El Faro de Motril	elfaromotril.es
El Independiente de Granada	elindependientedegranada.es
El Mira	elmira.es
El Noticiero Digital	elnoticierodigital.com
El País – Edición Andalucía	elpais.com
El Periodico de Chiclana	elperiodicodechiclana.com
El Puerto Actualidad	elpuertoactualidad.es
El Rincón Habla	elrinconhabla.com
El Sol de Antequera	elsoldeantequera.com
El perioódico de Mairena	elperiodicodemairena.com
Europa Press Andalucía	europapress.es
Europa Sur	europasur.es
Fuengirola TV	fuengirolatv.com
Granada Digital	granadadigital.es
Granada Hoy	granadahoy.com
Granada es Noticia	granadaesnoticia.com
GranadaDigital	granadadigital.es
Hora Jaén	horajaen.com
Hora Sur	horasur.com
Huelva Buenas Noticias	huelvabuenasnoticias.com
Huelva Costa	huelvacosta.com
Huelva Hoy	huelvahoy.com
Huelva Información	huelvainformacion.es
Huelva Red	huelvared.com
Huelva TV	huelvatv.com
Huelva24	huelva24.com
HuelvaYa	huelvaya.es
Ideal (Ed. Granada, Jaén, Almerí-a)	ideal.es
Ideal Granada	ideal.es
Ideal de Jaén	ttps:
Info Axarquía	infoaxarquia.es

Cofinanciado por la Unión Europea

Info Costa Tropical	infocostatropical.com
Info Linares	infolinares.com
Información Puente Genil	informacionpuentegenil.es
Información San Fernando	informacionsanfernando.es
Interalmería TV	interalmeria.tv
Jaén Hoy	jaenhoy.es
La Contra de Jaén	lacontradejaen.com
La Gaceta de Almería	lagacetadealmeria.com
La Guía del Sur	laguiadelsur.es
La Higuerita	lahiguerita.es
La Noción	lanocion.es
La Noción (Andalucía / Málaga)	lanocion.es
La Opinión de Málaga	laopiniondemalaga.es
La Razón	larazon.es
La Revista de Carmona	larevistacarmona.es
La Semana	periodicolasemana.es
La Voz Digital	lavozdigital.es
La Voz de Alcalá	lavozdealcala.com
La Voz de Almería	lavozdealmeria.com
La Voz de Morón	lavozdemoron.es
La Voz de Málaga	lavozdemalaga.com
La Voz de la Subbética	lavozdelasubbetica.es
La Voz del Sur	lavozdelsur.es
La Voz del Sur (ediciones Cádiz, Jerez, Sevilla, Málaga, Huelva, Granada, Córdoba, Almería, Jaén...)	lavozdelsur.es
Las 4 Esquinas	las4esquinas.com
Lebrija Digital	lebrijadigital.es
Linares 28	linares28.es
Lucena Digital	lucenadigital.com
Lucena Hoy	lucenahoy.com
Malaga Hoy	malagahoy.es

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Manilva TV	rtvmanilva.com
Marbella 24 Horas	marbella24horas.es
Marbella Confidencial	marbellaconfidencial.es
Marbella Directo	marbelladirecto.com
Mijas Comunicación	mijascomunicacion.org
Montilla Abierta	montillabierta.es
Montilla Digital	montilladigital.com
Morón Información	moroninformacion.es
Motril Digital	motrildigital.news
Málaga Actualidad	malagactualidad.es
Málaga Digital	diariomalagadigital.es
Málaga Es	malagaes.com
Málaga Hoy	malagahoy.es
Málaga al Día	malagaldia.es
Noticias 24 Digital	noticias24digital.com
Noticias Aljarafe	noticiasaljarafe.es
Noticias Motril	noticiasmotril.com
Noticias de Alcalá	noticiasdealcala.info
Noticias de Almeria	noticiasdealmeria.com
Noticias de Almería	noticiasdealmeria.com
Noticias de la Villa	noticiasdelavilla.net
OLA TV - EMA-RTV	ondalocaldeandalucia.es
Olé Benalmádena	olebenalmadena.com
Onda Algeciras TV	ondaalgecirastv.com
Onda Cero Andalucía	ondacero.es
Onda Cádiz TV	ondacadiz.es
Onda Jerez TV	ondajerez.com
Onda Local de Andalucía (Red EMA-RTV)	ondalocaldeandalucia.es
Onda Mencía (Doña Mencía)	ondamenciaradio.com
PTV Telecom (Córdoba, Málaga, Sevilla, Granada, Almería)	ptvtelecom.com

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Periódico de Málaga	periodicodemalaga.es
Portal de Cádiz	portaldecadiz.com
Puente Chico	puentechico.com
Puente Genil OK	puentegenilok.es
Puerto RealHoy	puertorealhoy.es
RTV Marbella	rtvmarbella.tv
Radio Almaina (Libre/Comunitaria)	radioalmaina.org
Radio Mijas	mijascomunicacion.com
Radio Polis (Comunitaria)	radiopolis.org.es
Radio Rute	radiorute.com
Ronda Semanal	rondasemanal.es
Rota al Día	rotaaldia.com
Sal Televisión	saltv.es
Sanlucar Información	sanlucarinformacion.es
Sevilla Actualidad	sevillaactualidad.com
Sevilla Buenas Noticias	sevillabuenasnoticias.com
Sevilla FC TV	sevillafc.es
Sur de Córdoba	surdecordoba.com
TG7 (Municipal)	tg7.es
TV Centro Andalucía	tvcentroandalucia.com
TVM Córdoba (Municipal)	tvm.cordoba.es
Tele Doñana	canaldonana.com
Telemotril	telemotril.com
Teleonuba	teleonuba.es
Teleprensa	teleprensa.com
Televisión Linares	facebook.com
Torremolinos Hoy	torremolinoshoy.com
Torremolinos TV	torremolinostv.com
UVitel TV	uvitelonline.com
Utrera Digital	utreradigital.com
Utrera Web	utreraweb.com

Cofinanciado por la Unión Europea

ReTo Red de Tolerancia contra los delitos de odio

Vega Digital	vegadigital.es
Videoluc TV	videoluc.com
Viva	vivacadiz.es
Viva Almería	vivaalmeria.es
Viva Almuñecar	vivaalmunecar.es
Viva Antequera	vivaantequera.es
Viva Arcos	vivaarcos.es
Viva Benalmádena	vivabenalmadena.es
Viva Costa del Sol Occidental	vivalacostaoccidental.es
Viva El Condado	vivaelcondado.es
Viva Estepona	vivaestepona.es
Viva Huelva	vivahuelva.es
Viva Jaén	vivajaen.es
Viva Marbella	vivamarbella.es
Viva Mijas	vivamijas.es
Viva Málaga	vivamalaga.net
Viva Sevilla	vivasevilla.es
Viva Torremolinos	vivatorremolinos.es
Viva Vélez Málaga	vivavelezmalaga.es
Víva Córdoba	vivacordoba.es
Víva Granada	vivagranada.es
elDiario.es Andalucía	eldiario.es
Área Costa del Sol	areacostadelsol.com
Écija Digital	ecijadigital.es
Écija Web	ecijaweb.com
Écija al Día	ecijaldia.es

Cofinanciado por la Unión Europea