

Escucha social sobre la violencia de género

Pablo Natal Gutiérrez, p.natal.2021@alumnos.urjc.es

Miguel Cuerdo Mir, miguel.cuerdo@urjc.es

Raquel Ibar Alonso, raquel.ibar@urjc.es

*Facultad de Ciencias Jurídicas y Sociales
Universidad Rey Juan Carlos*

RESUMEN

En esta investigación se procede a analizar publicaciones y conversaciones de la red social Twitter en torno a la temática de la violencia de género. De esta manera se elabora un análisis descriptivo de la volumetría de contenido en español y de su viralidad a través de las reacciones a los tweets: comentarios, "retweets" y "me gusta" producidos en los últimos años. Asimismo, se identifican tendencias, principales cuentas y palabras clave tanto de cuentas creadoras de contenido como de cuentas que reaccionan a la temática y se extraen otras categorías relacionadas y términos más comunes empleados para referirse a la violencia de género. Además, se analiza el sentimiento de las conversaciones por medio de procesamiento de lenguaje natural infiriendo las connotaciones positivas, negativas o neutras de los textos. A través de las cuentas citadas en los tweets, se procede a la identificación de personas o cuentas relevantes, tanto activas como pasivas, desde la perspectiva del género y localización. Estos atributos son inferidos a través de la información compartida en sus perfiles.

Área temática: A5 Aspectos cuantitativos de problemas económicos y empresariales

ABSTRACT

This research proceeds to analyse publications and conversations on the social network Twitter around the theme of gender violence. In this way, a descriptive analysis of the

volumetry of content in Spanish and its virality is elaborated through the reactions to the tweets: comments, "retweets" and "likes" produced in recent years. Likewise, trends, main accounts and keywords of both content creator accounts and accounts that react to the theme are identified, and other topics or common terms used to refer to gender violence are extracted. In addition, the sentiment of the conversations is calculated through natural language processing, inferring the positive, negative or neutral connotations of the texts. Through the accounts mentioned in the tweets, entities or relevant accounts are identified, both active and passive, considering gender and location inferred from information provided by the users on their profiles.

Palabras clave:

violencia de género; redes sociales; escucha social; twitter; NLP; procesamiento de lenguaje natural; análisis de sentimiento; análisis de género; inferencia de género; inferencia de localización

Área temática: A5 Aspectos cuantitativos de problemas económicos y empresariales.

1. INTRODUCCIÓN

Gracias a los avances tecnológicos, la era digital en la que vivimos actualmente, nos permite comunicarnos a través de múltiples canales en tiempo real interconectando personas sin necesidad de estar en el mismo lugar físico y, además, manteniendo un registro perenne y accesible en el que las publicaciones pueden perdurar en el tiempo como una hemeroteca del siglo XXI en formato digital.

En este nuevo paradigma de comunicación, surgen las redes sociales que han ido virando y encontrando diferentes propósitos para satisfacer las diversas necesidades de consumo de contenido por parte de las audiencias. En el caso de Twitter [1], una red social de microblogging fundada en 2006, el propósito es crear un espacio en el que los usuarios puedan publicar mensajes de texto breves, con una longitud inicial de hasta 140 caracteres que ha ido aumentando hasta los 280 caracteres actualmente. Esta red social permite a sus usuarios suscribirse o “seguir” otras cuentas con el objetivo de recibir las publicaciones de éstas, conocidas como tweets o tuits en castellano. Twitter cuenta en la actualidad con aproximadamente 400 millones de usuarios mensuales activos [2] lo que supone el doble de usuarios con los que contaba la red en el año 2012. Y las cifras siguen aumentando, ya que en el último año el crecimiento estimado ha sido de 44 millones de usuarios en todo el mundo y más de la mitad de estos clientes son activos visitando la red social a diario [3].

Twitter se posiciona en el número 9 del ranking mundial entre los sitios web más visitados [4] y se ha convertido en una plataforma de debate, divulgación, opinión, información y, en ocasiones, lo opuesto a ésta. La característica más relevante de esta red es el hecho de que todo ello ocurre en tiempo real y cualquier temática tiene cabida entre sus contenidos: actualidad, política, eventos, deporte o entretenimiento que se pueden seguir fácilmente por suscripción a los mismos desde el apartado “Temas”, o siguiendo a creadores de contenido afines.

Al igual que en los lugares físicos, en las plataformas virtuales surgen conversaciones e interacciones que, a diferencia del entorno presencial, pueden conectar a dos completos desconocidos que se encuentran por interés común hablando de un mismo tema.

En este contexto, el objetivo de este estudio es analizar y cuantificar volumétricamente el contenido en español que acontece en esta red social en torno a la violencia de género; tema abordado desde diversos puntos de vista según la tipología del usuario (medios de comunicación, políticos, celebridades, perfiles anónimos) que genera distintas reacciones y respuestas de los cibernautas.

2. ESTADO DE LA CUESTIÓN

2.1.1. Problemática de la violencia de género en países hispanohablantes

La violencia de género es una problemática que, entre los españoles, ha pasado de preocupar a un rango de población entre el 1%-2% en 2017, hasta el 6%-8% en los años previos al estallido de la crisis del COVID-19, según el CIS [5], siendo éste el último dato cronológico que proporcionó esta fuente cuyos resultados provienen de encuestas. A su vez, según datos del Ministerio de Igualdad [6], “el 57% de las mujeres confiesa haber sufrido algún tipo de violencia machista (física o sexual tanto en el ámbito de la pareja como fuera de él, acoso sexual o acoso reiterado o *stalking* a lo largo de sus vidas situando la cifra total en 11,7 millones de víctimas”.

Esto no sólo muestra el problema con el dato declarativo sino que, en una cifra medible y comportamental, según datos del Ministerio del Interior [7], alrededor de dos de cada diez detenidos en España lo son por delitos de malos tratos en el ámbito familiar, contra la libertad sexual o las relaciones familiares. Si atendemos a datos publicados por la Fiscalía General del Estado [8], en 2020 se han registrado un total de 150 785 denuncias, -0.89% respecto a 2019, cifra en crecimiento a lo largo de los años hasta el estallido de la crisis COVID-19, cuando se experimentó una reducción de hasta el 50% debido a las restricciones de movilidad. La serie de actuaciones de la Fiscalía por delitos sexuales se ha incrementado en los últimos años llegando a los 21 000 casos al año, cifra que duplica el promedio anual de años previos (2012 - 2015). Asimismo, se conoce que “los juzgados españoles recibieron un total de 42 035 denuncias por violencia de género en el cuarto trimestre de 2021, lo que supone un aumento del 13,03% respecto al mismo periodo del año anterior, según los últimos datos aportados por el Observatorio contra la Violencia Doméstica y de Género del Consejo General del

Poder Judicial (CGPJ). La cifra de asesinados por esta causa es de 1122 víctimas desde el 2003, año en el que comenzó el registro [9].

Este problema no sólo se circunscribe al territorio español. Según un informe recién publicado por la CEPAL, Comisión Económica para América Latina y el Caribe [10], “al menos 4091 mujeres fueron víctimas de feminicidio en 26 países de la región en 2020”. En el resto de los países que comparten la misma lengua hay estudios que muestran la alta prevalencia de la problemática, como por ejemplo en Perú donde “reportes oficiales apuntan que 7 de cada 10 mujeres adultas han sido objeto de violencia psicológica, física y/o sexual en algún momento de su vida” [11]; o en Colombia, donde 2 de cada 3 mujeres confiesan haberla sufrido, aunque sólo una de cada 10 lo denunciaría [12]. En el caso de Argentina, en el año 2015, el movimiento “*Ni una menos*” ayudó a visibilizar la problemática y estas iniciativas “constituyeron un punto de inflexión que evidenció la fuerza y masividad de la demanda por el tratamiento de dicha problemática” [13].

2.1.2. Estado del arte

La revisión de la literatura centrada en la temática de violencia de género identifica que existe una tendencia por parte de los gobiernos a ser más transparentes con la materia, como por ejemplo el Ministerio de Igualdad de España o el Ministerio de Justicia y Derechos Humanos de Argentina entre otros, que ofrecen públicamente acceso a los datos oficiales a través de informes y portales web específicos. En cuanto a la comunidad científica, se encuentran tanto estudios realizados por grupos de investigación del ámbito universitario como de institutos de investigación. En el contexto actual, hay análisis recientes enfocados a cuantificar el impacto de la crisis del COVID-19 sobre la violencia de género. En “*Violencias basadas en genero tiempos COVID-19*” [14] se concluye que, a pesar de una reducción de casos informados por canales físicos (entidades sanitarias y fuerzas del orden) dadas las restricciones de movilidad y confinamientos de la población, la pandemia mundial ha acelerado la digitalización, aumentando el número de contactos telemáticos y poniendo de manifiesto la necesidad de los gobiernos de implementar estrategias que permitan atender a las víctimas garantizando un servicio “accesible, de bajo costo y con

continuidad”. Adicionalmente, en el estudio “*Activismo y violencia de género en las redes sociales en la actualidad*” [15], en un contexto de creciente incidencia de la problemática en México, sentencia en sus conclusiones que “la red social es un medio eficaz para ayudar a combatir o disminuir la violencia de género”.

Existen estudios adicionales que, más allá de mostrar los datos y su evolución, aplican técnicas complejas o algoritmos de procesamiento natural de lenguaje (NLP) como los modelos de detección de actitudes peligrosas [16], del año 2017 y orientado a identificar usuarios que potencialmente puedan ser dañinos o perjudiciales a través de sus publicaciones en redes sociales. O más reciente, el ya citado [14], debido a la situación de pandemia mundial, donde se clasificaron textos en redes sociales encontrando otras temáticas relevantes y relacionadas con la violencia de género como el estrés financiero o la capacidad de contagio del virus y las respuestas internacionales.

En estas investigaciones los objetivos se circunscriben a un área geográfica o país en un periodo de tiempo concreto donde se obtienen unas conclusiones determinadas sobre la problemática de la violencia de género. En cuanto a otros estudios que abarquen con mayor amplitud histórica y que permitan evaluar su evolución, en un contexto homogéneo como es Twitter, no se han encontrado investigaciones de impacto publicadas hasta el momento.

2.1.3. Objetivos e hipótesis de la investigación

Considerando la oportunidad que nos brindan herramientas como la tecnología y la información pública disponible para extraer datos y analizarlos masivamente, el objetivo de este trabajo es extraer conclusiones que nos permitan entender con qué perspectiva, con qué palabras y cómo abordamos la temática de la violencia de género y qué términos, eventos, noticias y opiniones causan mayor repercusión social. La investigación se realiza tanto en los medios de comunicación como en la política que vivimos, tanto en España como en el resto de los países hispanohablantes con gran profundidad histórica.

De esta manera, se ha considerado ahondar más allá de lo presente con el objetivo de cruzar o contrastar datos históricos disponibles que representen la realidad de la problemática en cuanto a sucesos y hechos no declarativos, sino factuales, versus lo que se conversa en Twitter, una red social que nos permite disponer de información

representativa y fidedigna de las preocupaciones, opiniones y conversaciones de la población sobre la temática a estudiar. Con todo ello se consigue un método de análisis para descubrir la relevancia del asunto y cómo es tratado por esta población a estudio en base a sucesos, hechos o fechas concretas que elevan su notoriedad en comparación con el promedio u otros momentos desde el año 2010. Para cumplir el objetivo de investigación propuesto se plantean las siguientes hipótesis:

H1. Twitter no sólo es una red social sino una fuente significativa y fiable para la recogida de información sobre de violencia de género, permitiendo la extracción de aprendizajes e información valiosa que puede abrir nuevas líneas de investigación.

H2. Los datos nos indican una gran asimetría con mayor prevalencia de la mujer como víctima. Esto nos sugiere mayor contenido generado y vinculado a mujeres.

H3. Los diferentes movimientos sociales han repercutido en la sociedad, con la consiguiente mayor visibilidad y volumen de personas pronunciándose sobre el tema.

H4: Intensificar las conversaciones relacionadas con la violencia de género es un factor positivo que contribuye a hablar con mayor optimismo, sensibilizar a la población y, en última instancia, a reducir los casos.

Analizada la problemática y el estado de la cuestión que ha motivado esta investigación se procede a detallar su desarrollo partiendo de la metodología seguida, en la sección 3, y de todas las fuentes de datos utilizadas para el análisis en el punto cuatro. Por último, en los epígrafes 5 y 6, los resultados y las conclusiones obtenidas a partir de los datos y metodología mencionados anteriormente; además de futuras líneas de investigación que se podrían seguir tras la conclusión de esta investigación.

3. METODOLOGÍA

La metodología utilizada se basa fundamentalmente en la extracción tanto de todo el histórico de conversaciones sobre violencia de género como de los metadatos relacionados que existen en la red social Twitter en el periodo comprendido entre enero de 2010 hasta el mes completo más reciente disponible en el momento de la finalización de esta investigación, marzo de 2022. Esto supone más de 11 años completos de datos a estudiar a los que ha sido posible acceder gracias a la disponibilidad de la información y los permisos otorgados por Twitter Inc. para acceder como investigador. Esta licencia

concede el privilegio de acceso a un límite de 10 millones de consultas al mes. La gestión oficial de este permiso ha permitido obtener una cantidad masiva de referencias a las que aplicar las capacidades de computación con el objetivo de analizar un gran volumen de datos en un tiempo reducido y prácticamente en tiempo real.

Una de las funciones que ofrece Twitter que ha supuesto un factor relevante para esta investigación es la extracción de contenido bajo criterios de búsqueda avanzados [17] como la fecha, el idioma, o determinadas palabras clave. De esta manera para este estudio se ha utilizado el filtro de etiquetado que proporciona Twitter y que permite identificar automáticamente el idioma de un texto como contenido en español a través de sus algoritmos propietarios [18]. La red social permite hacer esa criba para no tener todo el volumen correspondiente a estas palabras clave, a pesar de que los hashtags puedan ser tanto palabras clave en inglés, como el movimiento *#MeToo*, como términos populares utilizados en España u otros países hispanohablantes.

3.1. Stack tecnológico

3.1.1. Entorno de desarrollo: Anaconda – Python 3

Para el desarrollo del proyecto se ha utilizado una distribución de software libre y de código abierto que soporta tanto código de Python como R. En concreto, para el presente análisis se ha utilizado la versión 3.8 de Python con las correspondientes librerías (Tabla 1) que han permitido desarrollar la investigación al completo, desde la fase inicial de extracción de los datos hasta el tratamiento de estos y la obtención de resultados y conclusiones.

Tabla 1. Librerías de Python

Librería	Descripción
csv	Importación y exportación de ficheros de texto plano tipo CSV.
datetime	Manejo de fechas y horas para el tratamiento de la información de este tipo.
json	Tratamiento y visualización de ficheros tipo JSON.
nltk	NLTK o Natural Language Toolkit permite realizar tareas de procesamiento natural de lenguaje (NLP).
os	Manejo del sistema operativo para el listado, creación, modificación y borrado tanto de ficheros como de carpetas.
pandas	Manipulación y análisis de datos en modo tabla (DataFrame) a partir de la lectura de ficheros de texto plano.
re	Tratamiento de textos con expresiones regulares para la limpieza de cadenas de caracteres.
requests	Conectividad, envío y tratamiento de peticiones HTTP (Hypertext Transfer

	Protocol) que permite emular tareas desde línea de comandos como el acceso a internet a través de un navegador web.
sqlite3	Interfaz de conectividad con bases de datos tipo PostgreSQL permitiendo crear, insertar y borrar en tablas, además de realizar consultas tipo SQL (structured query language o lenguaje de consultas estructurado).
string	Funciones de tratamiento de cadenas de texto: caracteres especiales, conversiones mayúsculas/minúsculas, etc..
time	Módulo para tratamiento de variable de tiempo.
twarc	Recolección y archivo de datos de Twitter en formato JSON vía API. Dispone de múltiples comandos implementados y el acceso a la nueva versión de la API (v2) para usuarios de tipo investigador (Academic Access).
tweepy	Acceso a la API de Twitter con múltiples funcionalidades pre-construidas.
wordcloud	Generador de gráfico de nube de palabras a partir pares entidades-número.

Fuente: Elaboración propia

3.1.2. Base de datos: SQL Lite3

La base de datos, de tipo PostgreSQL, cumple con la función de almacenamiento de la información descargada de Twitter (Figura 1) permitiendo la ejecución de consultas y la obtención de métricas por distintos niveles de agregación (usuarios, hashtags, etc.).

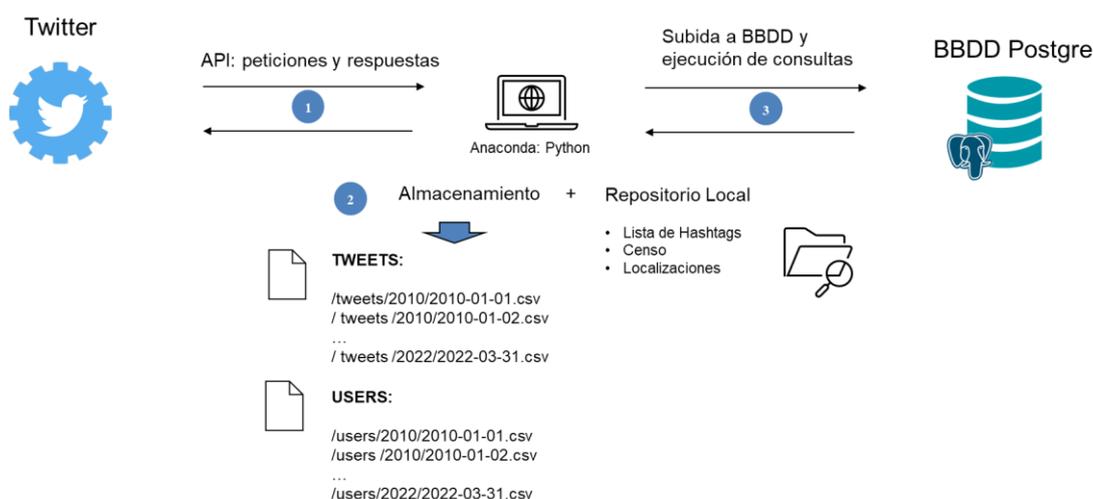


Figura 1.

Fuente: Elaboración propia

3.2. Extracción de datos de la red social Twitter

3.2.1. Tweets originales

La primera fase del proyecto consiste en la extracción de los datos de la red social Twitter. Para ello es necesario habilitar la API (*Application Programming*

Interfaces) o interfaz de aplicación programable con un usuario que sea capaz de conectarse por medio de código, en este caso desde el lenguaje de programación Python, de manera que se puedan realizar peticiones a la base de datos histórica de tweets. Según estudios de investigación [19][20], los tweets que incluyen hashtags reciben el doble de participación y, adicionalmente, incluir uno o más hashtags incrementa en un 55% la probabilidad de ser “retweeteado” y en un 21% más la participación con los tweets con más de dos hashtags. Por tanto, podemos afirmar que el hashtag es el elemento del que nace este estudio para poder diferenciar palabras clave dentro de la temática de la violencia de género.

El análisis comenzó con una primera extracción a partir de los hashtags más básicos relacionados con la temática: #machismo, #violenciadegenero o el popular e internacional #metoo. Una vez extraída toda la información de estos términos, con el objetivo de encontrar la mayor representatividad de toda la temática, se identificaron otros hashtags populares citados en los textos de los tweets. Por tanto, se procedió a actualizar la lista hasta obtener, finalmente, el listado definitivo (Tabla 2):

Tabla 2. Listado de hashtags

#25N	#feminazi	#laespañaquequieresesfeminista
#25Nov	#feminazis	#lamanada
#25Noviembre	#Femicidio	#LasFeministasQueremos
#4mujeresen48h	#femicidioemergencianacional	#lasperiodistasparamos
#8M	#feminism	#LaTecnologíaTambiénEsCosaDeNiñas
#8marzo	#Feminismo	#LeyLibertadSexual
#8mLaAlternativa	#FeminismoSindical	#luchafeminista
#asesinatosmachistas	#feminista	#Machismo
#BelieveSurvivors	#Feministas	#MachismoAsesina
#CifraDeLaVergüenza	#GeneracionIgualdad	#machismomata
#ComoMujerMeHaPasado	#girlpower	#marcha8m
#Cuéntalo	#GobiernoFeminista	#marchafeminista
#DiaContraLaViolenciaDeGenero	#HaySalida	#MeToo
#diadelamujer	#huelga8m	#miracomonosponemos
#diainternacionaldelamujer	#huelgafeminista	#miuteronoesnegocio
#ellassimerepresentan	#huelgafeminista2020	#MiVentreNoSeAlquila
#EllasValenOro	#huelgafeminista8m	#MujeresAutenticas
#EmergenciaFeminista	#HuerfanosVG	#mujeresenhuelga
#empoderamiento	#igualdad	#MujeresIgualesYLibres
#estaesnuestramanada	#igualdaddegenero	#mujeresyjusticia
#estamoshartasde	#InMujer	#MujerPorEncimaDeTodo
#EveryWoman	#internationalwomensday	
#felizdiadelamujer	#justiciapatriarcal	

#nadanospara8m	#nosqueremosvivas	#vientesdealquiler
#NiMasNiMenosIguales	#ObjetivoNiUnaMenos	#violenciacontralamujer
#NiUnaMenos	#pactodeestado	#ViolenciaDeGenero
#NiUnPasoAtras	#parodemujeres	#ViolenciaMachista
#NoAblacion	#parointernacionaldemujeres	#violenciasexual
#NocheVioleta	#patriarcado	#vivaslibresunidas
#noesabusoesviolacion	#PorEllas	#vivasnosqueremos
#NoEsNo	#SemanaDeLaMujer	#womensrights
#noestásola	#solosiessi	#yoparo
#NoHablesEnMiNombre	#sororidad	#YoSiTeCreo
#nolasdejesinvisibles	#SoyFeministaPorque	#yositecreohermana
#NoLasOlvidamos	#TerrorismoMachista	#yotambien
#nometoqueselutero	#TodasContamos	#yotecreo
#NoMiresAOtroLado	#TodasLasMujeres	#8M2021
#nonoscallamosmas	#TúTambién	#8Marzo2021
#NoSomosVasijas	#vagafeminista	#8deMarzo2021
#nosotrasparamos	#vagafeminista8m	
#nosparamos	#vagageneralfeminista	

Fuente: Elaboración propia

3.2.2. Información de cuentas.

A partir de los tweets originales se han conseguido los identificadores de todos los cibernautas presentes en Twitter que han mencionado algunos de los hashtags publicados previamente, de manera que se han extraído todos los metadatos de estos usuarios en relación a lo que ellos quieren dar a conocer dentro de la red social, pudiendo ser tanto su nombre o localización como una breve descripción de texto libre donde se encuentran intereses, motivaciones, creencias, inclinaciones políticas, etc.

3.3. Tratamiento y limpieza del dato

La metodología incluye una revisión de todo el juego de datos descargado de manera que su validación permita hacer una comparativa en igualdad de condiciones y evitando duplicados o ausencia de tweets en todo el periodo analizado. El objetivo consiste en disponer de una base de datos ordenada y homogénea para poder realizar diferentes consultas y extraer conclusiones.

3.3.1. Descarga y tratamiento de Tweets

Los tweets contienen tanto el texto publicado como los metadatos que permiten identificar al usuario que lo genera, cuándo se ha publicado, y las reacciones que ha

tenido ese contenido como el número de interacciones a través de ‘Me gusta’, ‘Retweets’ y comentarios. La información del identificador único del tweet sirve para evitar duplicados, además del hashtag que ha propiciado el encuentro del tweet.

Ha de tenerse en cuenta que esto ha ocurrido, es decir, al buscar tweets relacionados con la temática a través de diferentes palabras clave, se han obtenido tweets duplicados, resultado de mencionar múltiples hashtags, que han tenido que ser tratados a posteriori para ser eliminados y mantener una base de datos lo más representativa y aproximada a la realidad para las métricas analizadas. Aunque no es obligatorio ni siempre se utiliza un hashtag en cada publicación, como se mencionó en el epígrafe 3.2.1., su uso es muy extendido y aceptado en esta red social por proporcionar mayores probabilidades de relacionarse con una temática específica y generar más reacciones de otros usuarios con inquietudes similares. Además, sí que podemos asumir que, dada las volúmetrías, alrededor de 9,5 millones de tweets, la muestra es representativa para el objeto de análisis.

3.3.2. Obtención de hashtags y otras palabras clave

Con la totalidad de la base de datos dispuesta y accesible para su tratamiento y análisis no solo se dispone de los textos y una serie de valores informativos de volúmenes sino que se pueden extraer tanto los hashtags como otras palabras clave que se mencionan en todo el volumen disponible descargado de Twitter por medio del procesamiento del lenguaje natural (NLP), de manera que eliminamos palabras que no tienen valor semántico permitiéndonos filtrar para quedarnos solo con las que tienen interés y así poder detectar las volúmetrías de palabras más populares o entidades que están presentes en las conversaciones sobre violencia de género.

3.3.3. Análisis de sentimiento del texto del tweet

“Sentiment-spanish” es una biblioteca de Python que utiliza redes neuronales convolucionales para predecir el sentimiento de las oraciones en español. El modelo fue entrenado a partir de más de 800.000 reseñas de usuarios de las páginas ElTenedor, Decathlon, TripAdvisor, Filmaffinity y e-bay. Usando la tasa en las reseñas de los usuarios, se entrenó el modelo para que aprendiera del idioma con las librerías *Keras* y

Tensorflow, logrando una precisión de validación (precisión sobre datos nuevos, no utilizados para el entrenamiento) del 88%.

La puntuación del sentimiento de los diferentes tweets se ha obtenido con este algoritmo, por medio del lenguaje de programación Python, proporcionando un valor de estimación del sentimiento desde 0, como valor más negativo, hasta 1 como valor más positivo, pudiendo ser neutros en un punto intermedio.

3.3.4. Usuarios

De los usuarios de Twitter se obtienen datos directos a través del identificador numérico único del que todo usuario dispone; al igual que existe otro identificar único de tipo carácter donde, obtenido en el momento de registro y elegido por el usuario si está disponible, sólo está permitido utilizar caracteres alfanuméricos y el carácter ”_”.

La información disponible del usuario se obtiene referenciada al momento de la consulta ya que se provee del número de publicaciones hasta el momento en la red social, lo que permite cuantificar su presencia de manera activa y la interacción o la exposición que tienen a diversos contenidos dependiendo de las cuentas que siguen, con las que interactúan, además de su relevancia en base al número de seguidores que tiene. Otras métricas relevantes son los campos libres de texto en los que se puede informar dónde se localiza esa persona, o dice estar localizada, además de un breve resumen que pueda describir intereses u otros atributos personales como su unidad familiar y rol, profesión, aficiones o equipos deportivos a lo que apoya.

3.3.5. Localizaciones.

El cálculo de la ubicación inferido de los usuarios se basa en los campos disponibles “localización” y “descripción”. Aunque no es obligatorio, este atributo, de estar informado, siempre aparece visible de manera pública en Twitter. Este texto es libre, por lo que requiere un análisis de los distintos valores encontrados en los más de un millón de usuarios.

Para ello se han probado diferentes repositorios de datos y algoritmos de identificación de localizaciones en los que, pasado un texto como parámetro, atribuye tanto ciudades como países. La problemática de estos algoritmos, *locationtagger* [21] y *geotext* [22], radica en que están entrenados para dar mejor respuesta a textos en inglés

por lo que encuentra dificultades a la hora de clasificar términos en español; además de términos castellanizados para designar ciudades como Nueva York o Londres. Los textos de los que se dispone en este proyecto corresponden a cuentas de habla castellana, lo que refiere que la precisión del algoritmo en las distintas pruebas que se realizaron era muy baja, menor del 10% concretamente. Adicionalmente, el tiempo de computación era muy elevado. Por tanto, para enriquecer esta tasa de identificación, se ha construido una base de datos con municipios, ciudades y regiones de España basada en el censo del INE [23]; además de otras bases de datos con ciudades y países hispanohablantes [24].

3.3.6. Género.

Tras analizar diferentes soluciones para la inferencia del género, disponibles en inglés [26] y de pago [27] y con una baja tasa de éxito en español, a través de la extracción de tweets del usuario, se optó por elaborar un algoritmo de proximidad que contabiliza el número de veces que aparece en los diferentes y posibles campos (nombre del usuario, localización y descripción) alguno de los nombres de un censo de elaboración propia, es decir, se han tomado conjuntos de datos etiquetados con nombres de hombres y mujeres. Además de los datos del INE, se han encontrado listados de Argentina y Buenos Aires [28][29] con el censo de sus habitantes, etiquetado y clasificado diferenciando entre nombres de género masculino y femenino. Adicionalmente, se elaboró un listado con palabras clave inequívocas de género como las profesiones o rol familiar (“actor”, “actriz”, “papá”, “mamá”). Por último, se incluyeron términos identitarios de género binario y no binario como “él/he/him”, “ella/she/her”, “él/ella”. En el caso de que un nombre pueda existir para ambos géneros, no se ha asignado un género específico al usuario. Finalmente, se dispone de una base de datos de más de 64 000 nombres clasificados. Con esta información consolidada, se configuraron una serie de ficheros con el objetivo de encontrar estas diferentes entidades en los textos que se hallan en los campos citados anteriormente del usuario. La tasa final de reconocimiento de género asciende a 1.9 millones (28%).

3.4. Almacenamiento en base de datos y resultados

La información descargada y procesada se ha cargado en la base de datos mencionada. Este es el repositorio final que permite realizar consultas tipo SQL para la obtención de resultados.

4. DATOS

4.1. Datos privados: Twitter

4.1.1. Tweets

La recopilación de la información de los tweets se ha recogido por medio del uso de la API, en su versión 2.0. La documentación se encuentra disponible en línea. [30]

Información tanto del texto publicado como de otra información relevante y que es recogida a través del identificador único del tweet por medio del encontrado utilizando la API de Twitter que proporciona el texto del tweet junto con la marca de tiempo, el usuario y metadatos como los ‘Me gusta’, ‘Retweets’ y comentarios; además del hashtag o palabra clave que ha propiciado el encuentro del tweet.

Tabla 3. Identificador único

Nombre del campo	Tipo	Ejemplo	Descripción
Hashtag	Cadena	#8m	Identificador de la palabra clave en la búsqueda de tweets.
Username	Cadena	pablonatal	Nombre del usuario que publica el tweet.
User_id	Entero	320867504	Identificador numérico único del usuario que publica el tweet.
Id	Entero	8939392	Identificador numérico único del tweet.
Conversation_id	Entero	8939391	Identificador del tweet original. Si no es una respuesta, ambos campos (Id y conversation_id) son iguales.
Timestamp	Fecha y hora	2016-01-08 01:27:02	Fecha y marca de tiempo de publicación en huso horario CEST. (Central European Time (CET), UTC +1)
Replies	Entero	2	Número de comentarios publicados como respuesta al tweet original.
Retweets	Entero	10	Número de retweets que han reaccionado al tweet original, es decir, el usuario publica en su muro el tweet de un tercero.
Likes	Entero	12	Número de “Me gusta” que han reaccionado al tweet.
Is_retweet	Binario	1	Verdadero (1) o Falso (0)
Retweet_username	Entero	320864	Identificador del usuario que publica el tweet original si es un retweet. Vacío en caso contrario.
Retweed_id	Entero	8939234	Identificador del tweet original, si es un retweet. Vacío en caso contrario.

Text	Cadena	“Hola, esto es un tuit sobre el #8m”	Texto libre que publica el usuario y que ha de contener al menos un hashtag de los considerados en el alcance.
------	--------	--------------------------------------	--

Fuente: Elaboración propia

4.1.2. Datos de usuarios

La API de Twitter permite obtener información de los usuarios a través de un número de identificación único, extraído previamente con la descarga de cada tweet. La información disponible es: fecha de creación del perfil y actividad en la red.

4.1.3. Tablas calculadas

Se generan tablas calculadas intermedias para el análisis de resultados.

- Diarios: tabla con todos los tweets descargados.
- Únicos: tabla resultante de eliminar los tweets duplicados de la anterior.
- Hashtags: todos los hashtag buscados más los adicionales encontrados.
- Menciones: todos los usuarios pasivos o mencionados en un tweet.
- Sentimiento: todos los tweets con la puntuación de sentimiento.

4.2. Datos públicos

Por último, se encuentran disponibles numerosas fuentes públicas con informes y datos que ayudan a tener una visión más completa en cifras de la problemática.

Tabla 4. Fuentes de datos públicas

Variable	Fuente	Periodos
Víctimas mortales de violencia de género	Ministerio de Sanidad, Servicios Sociales e Igualdad. España	2003-2011 Anual
Denuncias de violencia de género	Observatorio contra la Violencia Doméstica y de Género del Consejo General del Poder Judicial (CGPJ). España	2009-2021 Trimestral
Denuncias de violencia de género presentadas directamente por la víctima en el juzgado	Consejo General del Poder Judicial (CGPJ). España	2009-2021 Trimestral
Denuncias de violencia de género presentadas por servicios asistencia a víctimas y terceros	Consejo General del Poder Judicial (CGPJ). España	2009-2021 Trimestral
Preocupación por la violencia contra la mujer (% población y multi respuesta)	Centro de Investigaciones Sociológicas. (CIS) España.	Abril 2010- Marzo 2020

Fuente: Elaboración propia

5. ANÁLISIS Y RESULTADOS

Los resultados obtenidos del análisis parten de una base de datos descargada de Twitter con más de 3,5GB de texto plano para analizar. En total, más de 9,45 millones de tweets generados por un total de 6,67 millones de usuarios produciendo las siguientes reacciones: 4,5; 40,4 y 71,6 millones de comentarios, ‘retweets’ y ‘Me gusta’.

5.1. Análisis descriptivo

5.1.1. Evolutivo

El análisis evolutivo (Figura 2) lleva a identificar cómo la temática alcanza máximos de actividad localizados en días muy específicos como el 25 de noviembre y el 8 de marzo que, además, se repiten año tras año. Si tenemos sólo en cuenta los días con más de 50 000 tweets, podemos apreciar cómo el 8 de marzo genera mayor volumetría mientras que el 25 de noviembre produce más interacciones como el número de tweets por usuario.

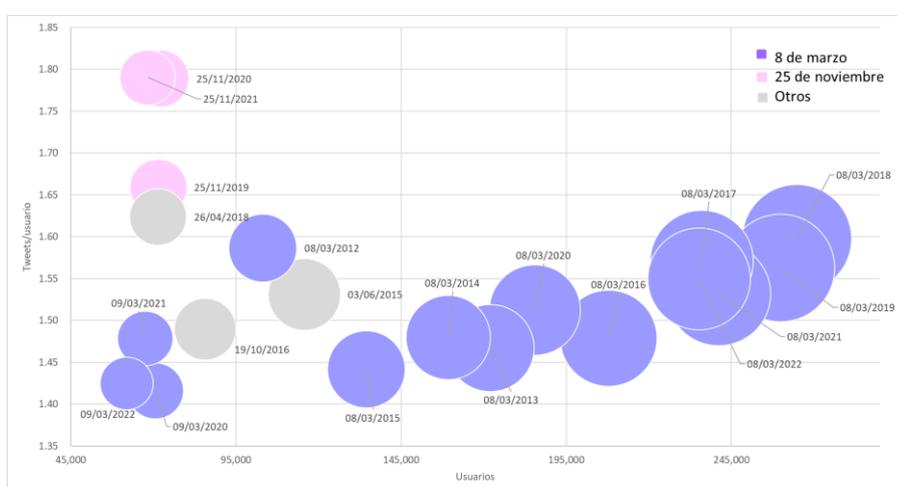


Figura 2. Análisis evolutivo de la temática

Fuente: Elaboración propia

Otros días destacados son:

- *3 de junio de 2015*: Manifestación en Argentina. Inicio del movimiento #NiUnaMenos [31].
- *19 de octubre de 2016*: “La muerte de Lucía, una adolescente que fue violada y asesinada en Argentina, ha llevado a una protesta masiva convocada en Argentina pero respaldada desde varias ciudades latinoamericanas en contra de la violencia de género” [32].
- *26 abril de 2018*: “La Audiencia Provincial de Navarra condena a los cinco acusados a nueve años de cárcel por abuso sexual continuado” [33], del caso de “La manada”.

Si nos enfocamos en el contenido por año y mes (Tabla 5), el mes que más contenido produce es marzo con el 39% y noviembre con el 12% consiguiendo más de la mitad de las conversaciones. En cuanto a los años, tal y como se veía en los datos del número de usuarios activos en la red social, el año 2018 obtuvo los máximos resultados históricos.

Tabla 5. Evolución del contenido

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dec	TOTAL
2010	211	422	4,784	850	814	895	1,913	2,375	3,376	3,331	5,240	3,603	27,814
2011	4,427	5,273	56,751	7,166	9,001	12,700	11,352	10,471	10,172	14,817	16,795	10,773	169,698
2012	9,093	10,712	137,148	10,795	12,384	11,080	9,828	12,788	14,428	23,242	141,031	11,990	404,519
2013	11,759	17,864	215,391	13,287	23,841	17,090	14,776	10,348	10,176	14,338	49,085	11,265	409,220
2014	10,759	12,591	219,679	25,734	50,644	11,131	10,393	20,044	14,101	14,675	51,733	13,803	455,287
2015	19,415	17,233	203,559	15,222	79,738	216,141	37,033	46,860	22,602	30,502	84,362	25,516	798,183
2016	23,005	48,154	279,548	38,409	39,151	87,426	33,253	77,434	29,639	194,692	102,539	39,062	992,312
2017	31,752	49,947	355,234	71,905	44,098	59,380	31,399	38,916	35,468	48,559	127,830	40,475	934,963
2018	52,626	77,217	448,997	215,617	92,148	108,399	52,843	45,921	43,431	60,794	111,596	180,367	1,489,956
2019	85,515	70,436	461,297	66,548	41,833	61,801	40,461	52,175	43,540	36,827	164,101	50,584	1,175,118
2020	41,981	77,763	427,922	42,744	44,281	70,818	44,554	46,613	58,734	45,035	143,758	40,064	1,048,267
2021	35,579	73,794	486,208	43,335	50,624	67,043	34,131	31,113	30,089	27,434	129,902	38,972	1,048,224
2022	24,404	33,152	400,190										457,746

Fuente: elaboración propia

5.1.2. Interacciones

Las diferentes métricas obtenidas de los tweets permiten ver tendencias en las interacciones. Mientras que el número de contenido generado por cada usuario decrece ligeramente, las respuestas, ‘retweets’ y comentarios a cada publicación ha crecido en los últimos años (Figura 3). El uso promedio de hashtags se ha calculado y demuestra una presencia masiva y habitual a partir de 2018 manteniendo un promedio entre 1 y 2 con mayor frecuencia en fechas señaladas de gran volumen de conversación.

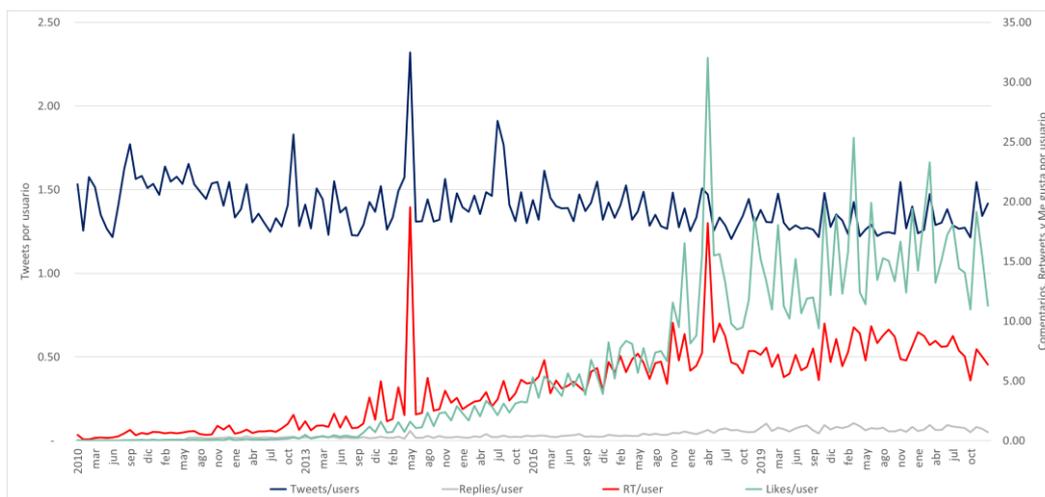


Figura 3. Tendencias de las interacciones

Fuente: elaboración propia

En cuanto a los hashtags, con los 22 más relevantes (Figura 4) se recoge el 80% del contenido destacado, además de los días de calendario ya mencionados, el movimiento #NiUnaMenos, #MeToo o el término #igualdad.



Figura 4. Principales hashtags

Fuente: elaboración propia

5.2. Género y entidades

Se han clasificado 1.9 millones de usuarios, de los cuales el 33% es identificable en género unívocamente con un 52.5% de hombres y 47.5% de mujeres. Ha de tenerse en cuenta que, según Statista [34] [35], se estima una presencia de hombres cercana al 68,5% a nivel mundial y al 62,8% en España.

5.2.1. Métricas: usuarios activos, 'retweeteados', con más 'Me gusta' y mencionados

Los volúmenes de publicaciones prueban mayor equilibrio por género. En cuanto a métricas de publicaciones aportan prácticamente de manera equitativa el 50% cada uno. Adicionalmente, las mujeres tienen mayor cuota de 'retweets' (55%) y 'Me gusta' (54%) mientras que reciben un ligero menor número de comentarios (48%) que los hombres.

5.2.2. Entidades

Entre las entidades más populares en las menciones de los tweets se encuentran políticos y partidos (PSOE, Ciudadanos, Pedro Sánchez, López Obrador); medios de comunicación y periodistas (El País, RTVE, El Diario, El Mundo, 20 minutos); entes públicos como Guardia Civil y Policía; y cuentas dedicadas específicamente a la temática (ONUMujeres, Feminicidio, InMujeres).

5.3. Análisis por países

De los 1,31 millones de usuarios para los que se ha identificado su país, del total de 6,67 millones, se sitúan en el top España, Perú y Argentina (Tabla 6).

Tabla 6. Ranking países

españa	22.32%
peru	22.06%
argentina	12.00%
colombia	8.49%
mexico	8.22%
venezuela	5.88%
chile	5.31%

Fuente: elaboración propia

Analizando el total de tweets por internautas, se puede concluir que los españoles son los que más tweets generan por usuario, seguidos por los mexicanos con el resto de los países de Latinoamérica en cifras muy similares. Analizando el resto de las métricas, destacar la menor tasa de ‘retweets’ de Colombia y de ‘Me gusta’ de Venezuela; también México destacando con más ‘Me gusta’ sobre el resto de los países.

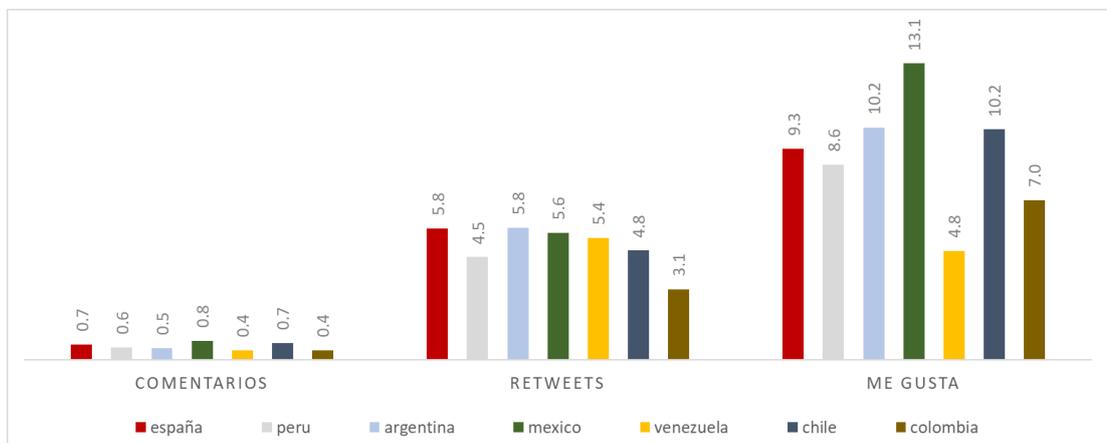


Figura 5.

Fuente: elaboración propia

5.3. Análisis de sentimiento

La evolución del sentimiento a lo largo del periodo analizado muestra una tendencia negativa aunque estable en los últimos años, en torno al 0.10, en un percentil muy bajo considerando el rango de 0 a 1. Como momentos estacionales positivos de

repunte se posicionan los meses de marzo, debido al movimiento 8M, y como momento negativo del año, diciembre-enero.



Figura 6.

Fuente: elaboración propia

5.3.1. Sentimiento por género

En cuanto a las diferencias en análisis de sentimiento entre hombres y mujeres no se han apreciado grandes diferencias destacables.

5.3.2. Sentimiento por país

La tendencia del sentimiento para todos los países se muestra muy similar con decrecimiento a partir de 2017; sólo hay una ligera mayor positividad por parte de España y México en el promedio total.

Tabla 7. Evolución del sentimiento

Países	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Promedio
argentina	0.17	0.18	0.18	0.18	0.18	0.19	0.18	0.17	0.13	0.14	0.13	0.12	0.13	0.16
chile	0.18	0.15	0.18	0.18	0.21	0.19	0.19	0.16	0.14	0.13	0.11	0.11	0.10	0.16
colombia	0.12	0.15	0.17	0.19	0.19	0.20	0.22	0.21	0.14	0.15	0.12	0.12	0.13	0.16
españa	0.18	0.18	0.21	0.21	0.22	0.21	0.20	0.20	0.15	0.14	0.13	0.13	0.14	0.18
méxico	0.16	0.19	0.18	0.22	0.21	0.22	0.22	0.21	0.15	0.14	0.12	0.12	0.13	0.18
perú	0.15	0.16	0.19	0.18	0.20	0.20	0.20	0.18	0.14	0.13	0.11	0.12	0.10	0.16
venezuela	0.14	0.13	0.15	0.17	0.22	0.21	0.22	0.20	0.15	0.15	0.11	0.12	0.11	0.16
Promedio	0.15	0.17	0.18	0.19	0.20	0.21	0.20	0.19	0.14	0.14	0.12	0.12	0.12	0.17

Fuente: elaboración propia

6. CONCLUSIONES

Los resultados obtenidos de esta investigación, dada la volumetría y la diversidad de contenido encontrado en relación a la temática de violencia de género, refrenda la hipótesis de que Twitter es un entorno digital en el que suceden muchas conversaciones entre personas permitiendo obtener métricas y resultados interesantes aplicables al ámbito social y muy valiosas para estudios e investigaciones.

Si bien es cierto que hay mayor presencia de hombres en esta red social se ha encontrado mayor equilibrio en la generación de contenidos e incluso mejores métricas de interacción ('retweets' y 'Me gusta') en perfiles femeninos validando la hipótesis de que las mujeres tienen mayor vinculación con la temática.

Como ya se anticipaba en el estado de la cuestión, las redes sociales ayudan a visibilizar problemas creando un entorno más accesible para hablar con mayor facilidad y menores restricciones. Según el CIS, [5], la preocupación por la temática de la violencia de género en el caso de España ha ido creciendo históricamente con mayor incidencia desde 2017 y, acorde a los resultados encontrados, existe mayor vinculación desde entonces a la hora de reaccionar a contenidos publicados.

Las cifras nos dicen que, aunque hay un ligero descenso en las víctimas mortales [9], la problemática sigue existiendo. De hecho, las denuncias han crecido en los últimos años y, aunque descienden las presentadas directamente por la víctima, las denuncias de violencia de género presentadas por terceros aumentaron significativamente, hasta el año 2018, para descender en años posteriores de manera muy similar a las volumetrías de tweets encontrados en este estudio. Esto nos permite concluir la importancia de las conversaciones que suceden en el ámbito de las redes sociales y la oportunidad de concienciar a la sociedad para colaborar y resolver de manera conjunta la problemática en el ámbito social.

7. REFERENCIAS BIBLIOGRÁFICAS

- [1] Página web oficial de la red social Twitter <https://twitter.com/>.
- [2] BRIAN DEAN (enero 2022). “How Many People Use Twitter in 2022?”. Blog de Blacklinko. <https://backlinko.com/twitter-users>
- [3] Statista (consultado en abril 2022). “Number of monetizable daily active Twitter users (mDAU) worldwide from 1st quarter 2017 to 4th quarter 2021”. <https://www.statista.com/statistics/970920/monetizable-daily-active-twitter-users-worldwide/>
- [4] CLAIRE BEVERIDGE. (marzo 2022). “Twitter Stats That Matter to Marketers in 2022”. Blog de Hootsuite. <https://blog.hootsuite.com/twitter-statistics/>
- [5] Centro de Investigaciones Sociológicas (marzo 2020). “Tres problemas principales que existen actualmente en España (Multirrespuesta %)”. https://www.cis.es/cis/export/sites/default/-Archivos/Indicadores/documentos_html/TresProblemas.html
- [6] Delegación del Gobierno contra la Violencia de Género. (2019) “Macroencuesta de violencia contra la mujer”. <https://violenciagenero.igualdad.gob.es/violenciaEnCifras/macroencuesta2015/Macroencuesta2019/home.htm>
- [7] Portal estadístico de criminalidad. “Detenciones e investigados por comunidades autónomas, tipología penal y periodo” (disponible hasta 2020). Ministerio del Interior. <https://estadisticasdecriminalidad.ses.mir.es/>
- [8] Memoria 2021 “Evolución de la criminalidad y actividad del Ministerio Fiscal”, Fiscalía General (2021). www.fiscal.es
- [9] “Estadística de Víctimas Mortales por Violencia de Género”. Observatorio contra la Violencia Doméstica y de Género del Consejo General del Poder Judicial (CGPJ). Ministerio de la Presidencia, Relaciones con las Cortes e Igualdad. (2022) <http://estadisticasviolenciagenero.igualdad.mpr.gob.es/>

- [10] L. ARROYO, A. BARRAGÁN, M. CENTENERA, I. SANTA EULALIA, N. GALARRAGA GORTÁZAR. CEPAL, (noviembre 2021) “Más de 4.000 feminicidios al año y mucha indignación: las razones de Latinoamérica para protestar el 25N”. El País. <https://elpais.com/sociedad/2021-11-25/mas-de-4000-femicidios-al-ano-y-mucha-indignacion-las-razones-de-latinoamerica-para-protestar-el-25n.html>
- [11] M. CONTRERAS-URBINA, G. GRANADOS DE ORBEGOSO, L. LEVANO (junio 2021), “Retos y alternativas para enfrentar la violencia de género en Perú”. <https://blogs.worldbank.org/es/latinamerica/retos-y-alternativas-para-enfrentar-la-violencia-de-genero-en-peru>
- [12] Infobae (noviembre 2021). “En Colombia fueron asesinadas 2,7 mujeres cada día en los últimos años; este es el panorama actual de la violencia de género en el país” <https://www.infobae.com/america/colombia/2021/11/25/en-colombia-fueron-asesinadas-27-mujeres-cada-dia-en-los-ultimos-anos-este-es-el-panorama-actual-de-la-violencia-de-genero-en-el-pais/>
- [13] P. CABRAL, J.A. ACACIO. (2016) “La violencia de género como problema público. Las movilizaciones por “Ni una menos” en la Argentina”. <https://ri.conicet.gov.ar/handle/11336/85429>
- [14] S. MARTÍNEZ RESTREPO, L. TAFUR MARÍN, J.G. OSIO, P. CORTÉS. (2020). “Violencias basadas en género en tiempos de covid-19”: análisis con Natural Language Processing”. <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/60696/IDL-60696.pdf>
- [15] Y.A. VITE HERNÁNDEZ, R. CORNELIO LANDERO, A. SUÁREZ OVANDO. (2020). “Activismo y violencia de género en las redes sociales en la actualidad”. <https://revistas.ujat.mx/index.php/perfiles/article/view/3903>
- [16] J. CALLONI, S. PAEZ, J. SALDARINI, J. CUEVAS, M. MULASSANO, A. BIANCIOTTI, E. SCARELLO, L. BANCHIO, F. DEGIOVANNI, L. SCHARFF (2017). “Modelo para Analizar Mensajes y Detectar Actitudes Peligrosas a través de Análisis de Sentimientos con Algoritmos de Aprendizajes”.

http://sedici.unlp.edu.ar/bitstream/handle/10915/61513/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y

- [17] Búsqueda avanzada de Twitter. <https://twitter.com/search-advanced?lang=es>
- [18] MITJIA T. (2015) “Evaluating language identification performance” Twitter Blog” https://blog.twitter.com/engineering/en_us/a/2015/evaluating-language-identification-performance
- [19] K. LEE. “How to Use Hashtags: How Many, Best Ones, and Where to Use Them” (2018). <https://buffer.com/library/a-scientific-guide-to-hashtags-which-ones-work-when-and-how-many/>
- [20] M.S LUCKIE (2012). “Best practices for journalists” Blog de Twitter. https://blog.twitter.com/official/en_us/a/2012/best-practices-for-journalists.html
- [21] Locationtagger (2020). Detect and extract locations (Countries, Regions/States & Cities) from text or URL. Also, find relationships among countries, regions & cities. <https://pypi.org/project/locationtagger/>
- [22] Geotext. MIT license. (2018). Geotext extracts country and city mentions from text. <https://pypi.org/project/geotext/>
- [23] “Demografía y población”. Instituto Nacional de Estadística. <https://www.ine.es>
- [24] “Sitio web con información de geografía, historia y deporte por país.” <https://www.sport-histoire.fr/es/>
- [26] “TwitterGenderPredictor. Python implementation of Sap et al.'s gender prediction algorithm for Twitter”. (2018). <https://github.com/jtwool/TwitterGenderPredictor>
- [27] “Namsor, name checker for gender, origin and ethnicity classification” <https://namsor.app/>
- [28] Registro Civil de nombres de Buenos Aires (consultado marzo 2022). <https://buenosaires.gob.ar/areas/registrocivil/nombres/busqueda/imprimir.php>

- Portal de nombre de datos del Gobierno de Argentina (consultado en marzo 2022). <https://nombres.datos.gob.ar/>
- [30] Documentación de la API de Twitter. “Developer Platform Documentation”. Twitter Inc. <https://developer.twitter.com/en/docs>
- [31] C. E. CUÉ, A.REBOSSIO. “Un multitud grita #NiUnaMenos en Argentina contra la violencia machista” (2015). El País. https://elpais.com/internacional/2015/06/03/actualidad/1433356172_949785.html
- [32] #NiUnaMenos: miles de mujeres se toman las calles de América Latina para exigir el fin de la violencia de género. (2016). BBC. <https://www.bbc.com/mundo/noticias-america-latina-37711297>
- [33] P.ORTEGA DOLZ. “La Manada se libra de la violación “ (2018), El País. https://elpais.com/politica/2018/04/26/actualidad/1524732098_409166.html
- [33] “Distribución porcentual de los usuarios de Twitter a nivel mundial en enero de 2021 por género” (2021) Statista. <https://es.statista.com/estadisticas/515756/twitter-distribucion-porcentual-de-los-usuarios-por-genero-en-espana/>
- [34] “Distribución porcentual de los usuarios de Twitter a nivel mundial en enero de 2021 por género” (2021) Statista. <https://es.statista.com/estadisticas/1203008/porcentaje-de-usuarios-de-twitter-por-genero/>