

Redes en Twitter y la defensa de la mujer peruana (ante el acoso y el feminicidio)

Patricia Daniela Reyes Silva

Máster Universitario en Ciencia de Datos

Área de Text mining & Social Network Analysis

Josep Maria Grau Masot

Jordi Casas-Roma

Junio 2019



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Redes en Twitter y la defensa de la mujer peruana (ante el acoso y el feminicidio)</i>
Nombre del autor:	<i>Patricia Reyes Silva</i>
Nombre del consultor/a:	<i>Josep Maria Grau Masot</i>
Nombre del PRA:	<i>Jordi Casas-Roma</i>
Fecha de entrega (mm/aaaa):	06/2019
Titulación:	<i>Máster Universitario en Ciencia de Datos</i>
Área del Trabajo Final:	Text mining & Social Network Analysis
Idioma del trabajo:	<i>Castellano</i>
Palabras clave	<i>Twitter, Big Data, Data Mining o Minería de Datos</i>

Resumen del Trabajo (máximo 250 palabras): *Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.*

Durante los últimos años, en Perú se observa un crecimiento importante en la tasa de casos de violencia contra la mujer y un número cada vez mayor de feminicidios. El objetivo de este Trabajo de final de Máster, es identificar a los distintos actores (organismos gubernamentales peruanos, colectivos privados, movimientos feministas, personalidades, etc.) y las comunidades que conforman, a la vez que se analiza que están hablando sobre este tema en redes sociales, específicamente, en Twitter.

La metodología utilizada ha contemplado:

- 1) La recolección de tweets (mediante etiquetas/hashtags y palabras clave circunscritos al territorio peruano), publicados entre el 4 y 12 de marzo, siendo el 8 de marzo del 2019, la fecha central en la que se celebra el día internacional de la mujer.
- 2) La aplicación de técnicas de Procesamiento del Lenguaje Natural, para analizar de una manera automática todos los mensajes, así como el uso de métricas de centralidad durante el análisis de redes, para la identificación de los perfiles más relevantes.
- 3) La conformación de la red de interacciones global en Perú sobre este tema, y el análisis del contenido de los retweets generados por los perfiles más relevantes.

Finalmente, el resultado de este trabajo ha permitido identificar los principales temas que las personas en la red discuten sobre la defensa de los derechos de las mujeres y que debieran tener principal atención por parte de las autoridades.

Abstract (in English, 250 words or less):

In recent years, Peru has seen a significant increase in the rate of cases of violence against women and an increasing number of femicides. The objective of this Final Master's Project is to identify the different actors (Peruvian government agencies, private collectives, feminist movements, personalities, etc.) and the communities that make up, while analyzing that they are talking about this issue in social networks, specifically, on Twitter.

The methodology used contemplates:

- 1) The collection of tweets (using tags / hashtags and keywords circumscribed to the Peruvian territory), published between March 4th and 12th, being March 8th 2019, the central date on which the international day of the woman.
- 2) The application of Natural Language Processing techniques, to automatically analyze all messages, as well as the use of centrality metrics during network analysis, to identify the most relevant profiles.
- 3) The creation of the global network of interactions in Peru on this subject, and the analysis of the content of the retweets generated by the most relevant profiles.

Finally, the result of this work has allowed us to identify the main issues that people in the network discuss about the defense of women's rights and that they should have primary attention from the authorities.

Índice

1. Introducción	1
1.1 Contexto y justificación del Trabajo	1
1.2 Objetivos del Trabajo	2
1.3 Enfoque y método seguido	3
1.3.1 Período de recolección de datos	3
1.3.2 Hashtags y palabras clave	3
1.4 Planificación del Trabajo	4
1.5 Breve sumario de productos obtenidos	4
1.6 Breve descripción de los otros capítulos de la memoria	4
2. Marco Teórico	6
2.1 Evolución de la Internet	7
2.1.1 Web 1.0	7
2.1.2 Web 2.0	7
2.1.3 Web móvil	8
2.2 Big Data	8
2.3 Big Data y las Redes Sociales	9
2.3.1 Surgimiento de las Redes Sociales	9
2.3.2 Breve cronología de las Redes Sociales en Internet	10
2.4 Las Redes Sociales en números	10
2.5 Twitter como Red Social	12
2.6 Twitter: Data Mining y Análisis de Redes	13
2.6.1 Terminología en Twitter	14
2.6.2 Datos que se pueden extraer desde Twitter	15
2.6.3 Redes de Retweets y métricas	16
3. Minería de Datos	21
3.1 Programación en Python	21
3.1.1 Programa para captura de tweets por fecha y palabra clave	21
3.1.2 Programa para lectura y proceso de archivos de tweets	22
3.1.3 Programa para creación de archivos para grafos	22
3.2 Filtrado y eliminación de datos	24
4. Modelado y análisis de las Redes	25
4.1 Red 8 de Marzo	28
4.1.1 Grafo de la red 8M del 4-7 marzo	29
4.1.3 Grafo de la red 8M del 9-12 marzo	41
4.2 Red Día de la Mujer	45
4.2.1 Resumen de pruebas estadísticas	45
4.2.2 Nodos influyentes por tipo de Centralidad	45
4.2.3 Visualización y exploración de la Red	46
4.2.4 Principales Comunidades	47
4.2.5 Nodos influyentes	47
4.2.6 Observaciones	48
4.3 Red Marcha 8M	50
4.3.1 Resumen de pruebas estadísticas	50
4.3.2 Nodos influyentes por tipo de Centralidad	50
4.3.3 Visualización y exploración de la Red	51

4.3.6 Observaciones.....	53
4.4 Red de Acoso	54
4.4.1 Resumen de pruebas estadísticas.....	54
4.4.2 Nodos influyentes por tipo de Centralidad	54
4.4.5 Nodos influyentes	56
4.4.6 Observaciones.....	57
4.5 Red de Femicidio.....	58
4.5.1 Resumen de pruebas estadísticas.....	58
4.5.2 Nodos influyentes por tipo de Centralidad	58
4.5.3 Visualización y exploración de la Red.....	59
4.5.4 Principales Comunidades.....	60
4.5.5 Nodos influyentes	60
4.5.6 Observaciones.....	61
5. Análisis y Observaciones	63
5.1 Análisis Estructural de las Redes	63
5.2 Análisis por Comunidades	64
5.3 Análisis Textual	67
6. Conclusiones.....	69
7. Glosario.....	72
8. Bibliografía.....	75
9. Anexos	77

Lista de figuras

Figura 1. Perú - Estadísticas de Femicidio 2011-2018	1
Figura 2. Digitalización alrededor del mundo en 2019	6
Figura 3. Crecimiento Digital Anual (Annual Digital Growth)	7
Figura 4. Las seis Vs del Big Data	9
Figura 5. Penetración de las Redes Sociales por Región	11
Figura 6. Usuarios activos mensualmente en Facebook	11
Figura 7. Visión general de la audiencia de Twitter	12
Figura 8. Gráfico de red dirigida básica.....	16
Figura 9. Estructura de archivo de Tweets (de tipo Json)	21
Figura 10. Estructura de datos de archivo de nodos (de tipo Csv)	22
Figura 11. Estructura de datos de archivo de interacciones (de tipo Csv)	23
Figura 12. Relación de archivos de nodos y enlaces (de tipo Csv)	24
Figura 13. Diseño Fruchterman-Reingold para la Red 8M del 4-7 marzo	32
Figura 14. Diseño Force Atlas 2 para la Red 8M del 4-7 marzo.....	33
Figura 15. Diseño Fruchterman-Reingold para la Red 8M del 8 marzo	37
Figura 16. Diseño Force Atlas 2 para la Red 8M del 8 marzo	38
Figura 17. Diseño Fruchterman-Reingold para la Red 8M del 9-12 marzo	42
Figura 18. Diseño Force Atlas 2 para la Red 8M del 9-12 marzo.....	42
Figura 19. Participación de políticos en la red 8M.....	44
Figura 20. Diseño Fruchterman-Reingold para la Red Día de la Mujer	46
Figura 21. Diseño Force Atlas 2 para la Red Día de la Mujer	46
Figura 22. Interacciones del MIMP en el Día Internacional de la Mujer	49
Figura 23. Diseño Fruchterman-Reingold para la Red Marcha 8M	51
Figura 24. Diseño Force Atlas 2 para la Red Marcha 8M.....	51
Figura 25. Diseño Fruchterman-Reingold para la Red Acoso en Perú.....	55
Figura 26. Diseño Force Atlas 2 para la Red Acoso en Perú	55
Figura 27. Diseño Fruchterman-Reingold para la Red Femicidio en Perú ...	59
Figura 28. Diseño Force Atlas 2 para la Red Femicidio en Perú	59
Figura 29. Enlaces de la Fiscalía Perú dentro de la Red de Femicidio	62
Figura 30. Enlaces del MIMP dentro de la Red de Femicidio	62
Figura 31. Tabla de participación de las Comunidades por Red	65
Figura 32. Word Cloud de la Red de Femicidio	67
Figura 33. Word Cloud de la Red de Acoso	68

Lista de tablas

Tabla 1. Red 8M del 4-7 marzo - Resumen de pruebas estadísticas.....	31
Tabla 2. Red 8M del 4-7 marzo - Nodos influyentes por tipo de Centralidad ...	32
Tabla 3. Red 8M del 4-7 marzo - Principales Comunidades	34
Tabla 4. Red 8M del 4-7 marzo - Nodos influyentes	35
Tabla 5. Red 8M el 8 de marzo - Resumen de pruebas estadísticas	36
Tabla 6. Red 8M el 8 de marzo - Nodos influyentes por tipo de Centralidad ...	36
Tabla 7. Red 8M el 8 de marzo - Principales Comunidades	39
Tabla 8. Red 8M el 8 de marzo - Nodos influyentes.....	40
Tabla 9. Red 8M del 9-12 marzo - Resumen de pruebas estadísticas.....	41
Tabla 10. Red 8M del 9-12 marzo - Nodos influyentes por tipo de Centralidad	41
Tabla 11. Red 8M del 9-12 marzo - Principales Comunidades	43
Tabla 12. Red 8M del 9-12 marzo - Nodos influyentes	44
Tabla 13. Red Día de la Mujer - Resumen de pruebas estadísticas	45
Tabla 14. Red Día de la Mujer - Nodos influyentes por tipo de Centralidad.....	45
Tabla 15. Red Día de la Mujer - Principales Comunidades.....	47
Tabla 16. Red Día de la Mujer - Nodos influyentes	48
Tabla 17. Red Marcha 8M - Resumen de pruebas estadísticas.....	50
Tabla 18. Red Marcha 8M - Nodos influyentes por tipo de Centralidad	50
Tabla 19. Red Marcha 8M - Principales Comunidades	52
Tabla 20. Red Marcha 8M - Nodos influyentes	53
Tabla 21. Red de Acoso - Resumen de pruebas estadísticas.....	54
Tabla 22. Red de Acoso - Nodos influyentes por tipo de Centralidad	54
Tabla 23. Red de Acoso - Principales Comunidades	56
Tabla 24. Red de Acoso - Nodos influyentes	57
Tabla 25. Red de Femicidio - Resumen de pruebas estadísticas.....	58
Tabla 26. Red de Femicidio - Nodos influyentes por tipo de Centralidad	58
Tabla 27. Red de Femicidio - Principales Comunidades	60
Tabla 28. Red de Femicidio - Nodos influyentes	61
Tabla 29. Escala para la medición de la Densidad de Red	63
Tabla 30. Categorización de la densidad por Red.....	64
Tabla 31. Modularidad y Comunidades por Red	64

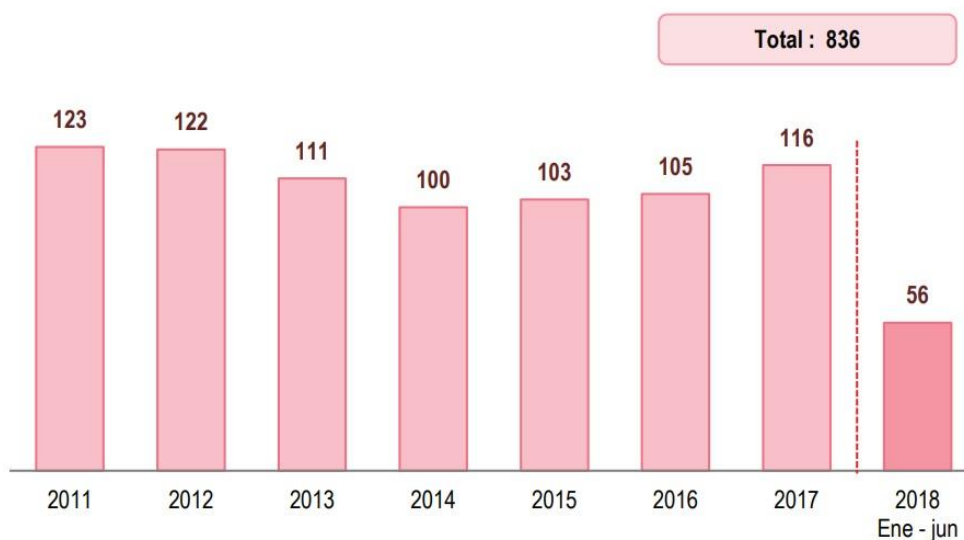
1. Introducción

1.1 Contexto y justificación del Trabajo

En el Perú, se vive actualmente una crisis de inseguridad ciudadana, y somos las mujeres uno de los sectores de la población más afectado.

Según datos del Instituto Nacional de Estadística e Informática (INEI), en su publicación: Perú - Estadísticas de Femicidio 2011-2018, la cifra de feminicidios asciende a 836 víctimas hasta junio del 2018, siendo las mujeres jóvenes las más vulnerables a este tipo de homicidio¹ [1]

PERÚ: NÚMERO DE VÍCTIMAS DE FEMINICIDIO, 2011 - 2017 Y ENERO - JUNIO 2018



Nota: Fecha de corte 30/06/2018.

Fuente: Ministerio Público - Observatorio de Criminalidad.

Elaboración: Instituto Nacional de Estadística e Informática.

Figura 1. Perú - Estadísticas de Femicidio 2011-2018

Como se observa en la figura 1, desde enero del 2011 al 30 de junio del 2018, el Femicidio ha cobrado 836 víctimas. El año 2018 concluyó con un total de 149 feminicidios, la cifra anual más alta desde el 2011 y en los dos primeros meses del 2019 ya se han registrado 22 casos.²

Esta dura realidad se debate a diario en las redes sociales en Perú, y es Twitter una de las redes donde las opiniones y sentimientos de muchos peruanos se expresan sobre este delicado tema, situación que se ha convertido en un problema de índole nacional.

¹ [Perú: Estadísticas de Femicidio 2011-2018](#)

² Información recopilada por [América Noticias](#), según la base de datos y cifras del [Ministerio de la Mujer y Poblaciones Vulnerables](#)

En mi opinión particular, pienso que para resolver este problema, las autoridades competentes (Ministerio de la Mujer y Poblaciones Vulnerables, otros ministerios, Fiscalía, Policía, Poder Judicial, la Defensoría del Pueblo, las municipalidades, etc.) deben trabajar articuladamente a fin de fortalecer los sistemas de protección a la mujer, a la vez que las mujeres agredidas “no deben callar”, sino por el contrario, “deben denunciar a sus agresores” y simultáneamente, se deben implementar campañas de educación y respeto de género durante todo el período escolar/universitario a nivel nacional.

Este trabajo no pretende dar solución a este grave problema de agresión a la mujer peruana, pero aporta en el análisis de esta situación en cuanto a su repercusión y discusión en las redes sociales (ya que si no existiera este problema, nadie hablaría de él).

Finalmente, como resultado de este trabajo, se conocerá a los principales actores que, desde distintas comunidades, participan en pro de los derechos de la mujer peruana (y aquellos otros, que debiendo figurar, tal vez no lo hacen). Este análisis también podría ser utilizado por las autoridades competentes involucradas en la defensa de la mujer, para conocer y evaluar su participación en las redes sociales, específicamente en Twitter, a fin de redirigir, reorientar y dar a conocer sus planes de acción en favor de los derechos de la mujer peruana.

1.2 Objetivos del Trabajo

Teniendo como marco la celebración del Día Internacional de la Mujer, el pasado 8 de Marzo, se ha establecido los siguientes objetivos:

- Modelar y diseñar las principales redes en Twitter, que conformadas por los principales actores en Perú, se identifican con la defensa de los derechos de la mujer.
- Identificar a los principales actores o influyentes dentro de cada red, así como las comunidades a las que pertenecen.
- Descubrir y describir las principales relaciones entre los principales actores y sus comunidades.
- Identificar si existen relaciones o no entre actores del Gobierno y de la sociedad civil, ambos interesados por los derechos de las mujer.

1.3 Enfoque y método seguido

Durante la etapa de minería de datos, se realizaron pruebas iniciales de extracción de datos (tweets) circunscritos en el área de geo-localización de Perú (-84.64,-20.2,-68.65,-0.04) y particularmente de la ciudad de Lima (-77.1992,-12.5199,-76.6208,-11.5724). A partir del resultado obtenido, se comprobó que son escasos los usuarios peruanos que en Twitter tienen activada la opción de localización, por ende, los datos capturados fueron insuficientes para el análisis.

Ante esta situación, y para conseguir los objetivos propuestos, se optó por recolectar los tweets a través de 'hashtags' y palabras clave en la zona geográfica de Perú y 2500km alrededor (longitud -9.189967 y latitud -75.015152).

1.3.1 Período de recolección de datos

Los datos fueron capturados del 4 al 12 de Marzo, teniendo como fecha central el 8 de Marzo del 2019, día en el que se celebró a nivel mundial el "Día Internacional de la Mujer".

1.3.2 Hashtags y palabras clave

Las siguientes fueron las etiquetas o palabras clave utilizadas para la extracción de tweets:

#8M
#Marcha8M
#DiaInternacionalDeLaMujer
#DiaDeLaMujer
#8marzo2019
#8M2019
Feminicidio
Acoso
Feminismo
Feministas

A partir de una gran cantidad de datos capturados, se aplicaron técnicas de minería de texto para descartar los tweets no pertenecientes al área geográfica (longitud/latitud) y locación correspondiente a Perú, calculando luego métricas de análisis de redes para la identificación de las comunidades y principales actores.

1.4 Planificación del Trabajo

En el Anexo 1, se visualiza mediante un diagrama de Gantt, la programación de las tareas ejecutadas durante el desarrollo del presente trabajo de fin de Máster, desde la selección del tema hasta su defensa.

Se puede observar que en cada etapa del proyecto, se interactúa con el asesor, con el cual se busca validar los avances y recibir recomendaciones que fueron discutidas y tomadas en cuenta previamente a la presentación de cada Prueba de Evaluación Continua (PEC).

Los hitos importantes a lograr se encuentran identificados por las tareas denominadas “Entregable #PEC Completado”, y es importante también observar, que la tarea que mayor tiempo de ejecución presentó fue la de Minería de Datos (I y II), la que demandó un total de 21 días para su desarrollo.

En lo relacionado al recurso tiempo utilizado, se programó una inversión de 4 horas diarias de lunes a viernes y de 6 horas los fines de semana.

1.5 Breve resumen de productos obtenidos

El presente trabajo de fin de Máster tendrá como principales entregables:

- El análisis de cinco redes en Twitter, identificando a los principales actores peruanos que se relacionan con la defensa de los derechos de la mujer.
- La visualización de dos nubes de palabras (Word Clouds), desarrollados a partir de los retweets de los principales actores influyentes de la red de Acoso y de Femicidio.
- Un resumen de las observaciones y conclusiones alcanzadas como parte final del presente documento, así como las recomendaciones para trabajos futuros.

1.6 Breve descripción de los otros capítulos de la memoria

La memoria del trabajo final de máster se estructura de la siguiente forma:

- **Capítulo 1: Introducción.** Contiene la introducción, contextualización, justificación y objetivos del presente trabajo, detallando así mismo, la metodología o enfoque seguido según un plan trabajo que se ajusta a los requerimientos u objetivos planteados.
- **Capítulo 2: Marco Teórico.** Contiene el desarrollo del análisis del mercado (estado del arte), que intenta justificar el presente trabajo de fin de Máster, a través de la teoría de otros trabajos o investigaciones.

- **Capítulo 3: Minería de Datos.** Contiene la descripción del proceso de extracción y pre-procesado de los datos utilizados en este trabajo. Los datos se extraen desde Twitter utilizando una biblioteca de Python llamada Tweepy (API de Twitter), y circunscribiendo el área de análisis a usuarios de Twitter localizados en Perú.
- **Capítulo 4: Modelado y análisis de las Redes.** Contiene las consideraciones para el análisis y creación de las redes (basadas en Retweets), la cuales estarán representadas por gráficos dirigidos. Así mismo se calculan las principales estadísticas y se identifican a los diez usuarios influyentes por cada red, según las medidas de centralidad de grado, de intermediación y de vector propio. Se visualizan también las redes, de forma general, siguiendo los algoritmos Force Atlas 2 y Fruchterman-Reingold.
- **Capítulo 5: Análisis y Observaciones.** Contiene la descripción, análisis y observaciones de las principales comunidades detectadas en cada red a partir de la función de "Modularidad" dentro la plataforma Gephi, así como el detalle de las gráficas a nivel de las comunidades, es decir, la interpretación de cuán "activo" y qué tan conectado estuvo el usuario con los otros.
- **Capítulo 6: Conclusiones.** Contiene las conclusiones resultantes al finalizar este trabajo, y las recomendaciones a futuro.
- **Capítulo 7: Glosario.** Contiene la descripción de algunos términos usados.
- **Capítulo 8:** Bibliografía utilizada en la presente memoria.
- **Capítulo 9:** Anexos

Así mismo, otro dato interesante observado en la figura 3, es que en el último año, el número de usuarios de Internet y de las redes sociales han aumentado en más del 9%.



Figura 3. Crecimiento Digital Anual (Annual Digital Growth)

2.1 Evolución de la Internet

El uso de la Internet ha evolucionado desde la incorporación de la World Wide Web (WWW), allá en los años 90 hasta el día de hoy. Hemos pasado de contar con 50 sitios web que existían en el mundo a finales de 1992 y de 150 en 1993[3] a más de 1.000 millones de sitios en Internet (1.518.207.412 según la encuesta de servidores web de de enero de 2019 en comparación con 1.805.260.010 en enero de 2018). [4]

Así mismo, la tecnología utilizada en Internet fue mejorando debido a la mayor demanda de servicios y nuevos usos, pudiéndose identificar tres etapas de desarrollo [5]:

2.1.1 Web 1.0

Cuando se utilizaba la Internet principalmente para listas de correo, fotos, boletines, blogs personales y sitios web estáticos, en lugar de HTML dinámico. La velocidad de acceso era baja y los sistemas de almacenamiento de archivos se incrementaban cada vez más.

2.1.2 Web 2.0

Cuando se introdujo la interoperabilidad y la usabilidad en los sitios web, el usuario pudo entonces agregar contenido (sitios web dinámicos) y ello conllevó al desarrollo y uso de las redes sociales, plataformas de vídeos y aplicaciones

web. Las aplicaciones de escritorio empezaron entonces a migrar y poblar la Internet.

2.1.3 Web móvil

Una vez que los teléfonos móviles se conectaron a Internet, el contenido de los sitios web fue adaptado de la versión de escritorio a la de los smartphones, abriendo el camino al auge las redes sociales, plataformas de vídeos y aplicaciones web que hoy en día utilizamos.

Lo común y constante en estas etapas, fue la progresiva incorporación e intercambio de datos en la red. La era de los “grandes datos”, o Big Data (por su traducción en inglés), ya tenía lugar, y tanto los sistemas de almacenamiento y acceso a los datos se adaptaron para soportar y dar respuesta a los requerimientos de los usuarios en tiempos razonables.

2.2 Big Data

Hoy en día, es común que compartamos nuestra vida en la Internet cuando realizamos publicaciones en línea, usamos texto, audio, fotos y video, en un contexto de Big Data. Los ingentes cantidades de datos (big data), pueden provenir de diferentes e innumerables fuentes, tales como: sistemas de transacciones comerciales, bases de datos de clientes, registros médicos, registros de flujo de clics de Internet, aplicaciones móviles, redes sociales, los resultados recopilados de experimentos científicos, datos generados por máquinas y sensores de datos en tiempo real [6], etc. y dichos datos pueden seguir una estructura fija, una estructura variable (o semi-estructurada), o lo más probable: no contar con alguna estructura conocida. En cualquiera de los casos, son muchos los datos que se almacenan y que de ser requerido para su análisis deben ser pre-procesados utilizando herramientas de extracción de datos o software de preparación de datos [6].

Es conocido que Big Data se caracteriza generalmente por las tres “V”:

- El **V**olumen extremo de datos,
- La gran **V**ariación de tipos de datos y,
- La **V**elocidad a la que se debe procesar dichos datos.

Esas características fueron definidas por primera vez por Doug Laney, analista de Gartner, en un informe publicado en 2001 [7], pero luego se fueron agregando otras “V” a la descripción inicial de Big Data, tales como: la veracidad, el valor y la variabilidad como se visualiza en la figura 4.

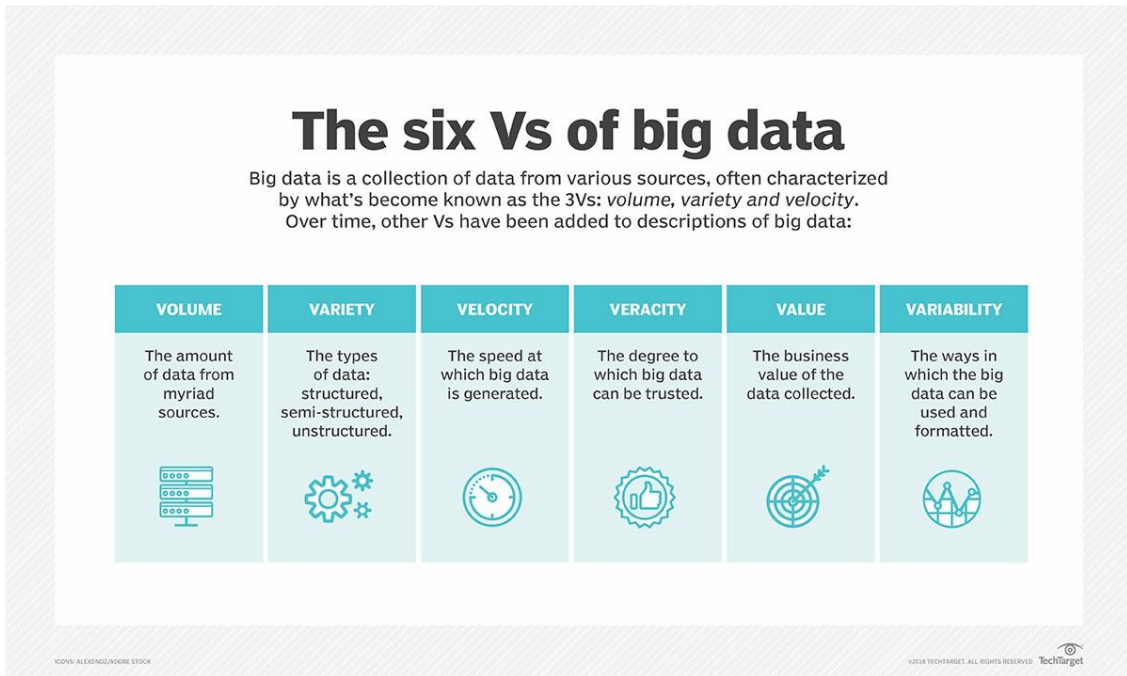


Figura 4. Las seis Vs del Big Data

Es interesante señalar que existen otros aportes a la redefinición del volumen de Big Data, como el de Ripon Patgiri, quien señala que se encuentran comprometidas otras V (voluminosidad, vacío y vitalidad) y añade dos V más al paradigma del Big Data: vendedor y vase (florero). [8]

Otro dato importante es, que si bien el Big Data no equivale a ningún volumen específico de datos, el término se usa a menudo para describir terabytes, petabytes e incluso exabytes de datos capturados a lo largo del tiempo. [6]

2.3 Big Data y las Redes Sociales

2.3.1 Surgimiento de las Redes Sociales

Los humanos somos seres sociales, necesitamos interactuar con otros humanos durante el transcurso de nuestras vidas, y ello nos hace desarrollar en todo aspecto, desde lo personal hasta lo profesional.

Revisando en breve la historia de la humanidad, el deseo del hombre por comunicarse casi instantáneamente data desde el año 500 A.C cuando utilizaban medios como el caballo, y en América el uso de un sistema de comunicaciones del Imperio Inca denominado **Chasqui** (del quechua: chaski/postillón o también chaski: correo, persona de relevo), constituido por jóvenes mensajeros que recorrían “corriendo” los caminos Incas para llevar un mensaje o recado. [9] (las señales de humo no han sido consideradas por tratarse de comunicaciones de índole visual).

Más tarde, con la llegada del telégrafo en 1792 las comunicaciones se aceleraron más que con el caballo, y en 1865 con el desarrollo del neumático las distancias se acortaron, pero no fue sino hasta inicios de siglo XIX cuando todo cambió con la invención del teléfono y la radio, que sumadas a la

televisión posteriormente y la Internet más recientemente, constituyen hasta la actualidad, los más importantes medios de redes sociales.

Como vemos, las redes sociales han existido casi desde siempre, y están constituidas por un conjunto de personas que se conectan y relacionan para intercambiar información de interés común. Independientemente de la ubicación geográfica, y el medio utilizado para la comunicación, dichas redes han evolucionado para reducir distancias y ser más veloces, así, en nuestro tiempo actual, podemos observar que las redes sociales se valen de aplicaciones y sitios web para comunicar e interactuar a través de la Internet.

2.3.2 Breve cronología de las Redes Sociales en Internet

En 1997, surge la primera red social online denominada Sixdegrees.com, que permitía colocar una foto de perfil y conectar con amigos a través de mensajes hasta alcanzar seis niveles de conexión o afinidad, por ello su nombre. [10]

En 1998, aparece Moveon.org, como uno de los primeros sitios activistas opuestas al juicio político de Bill Clinton, en 1999 surgen los sitios de blogs que comienzan a ser populares, así como sitios web comerciales financiados con publicidad (2000). Desde entonces, las redes sociales online se nutren de innumerables sitios: Wikipedia (2001), Friendster (2002), LastFM, uno de los primeros sitios de música en línea en 2002, LinkedIn, MySpace, WordPress, Photobucket y Flickr (2003), la versión de Facebook de Harvard, Care2, Multiply, Ning, Orkut, Mixi, Piczo y Hyves (2004), Youtube, Yahoo! 360, Bebo y Reddit (2005), Facebook y Twitter (2006), Tumbler y Friendfeed (2007), Spotify, Ping, Groupon y Kontain (2008), Foursquare (2009), Instagram, Pinterest y Google Buzz (2010), Google+ (2011).

Del 2011 al 2018, las redes sociales constituyen un ecosistema virtual en el que convivimos individuos y empresas, que si bien, su futuro es incierto [10], de lo que si se tiene certeza, es que cada día, el volumen de datos que se comparte por dichas redes sigue en aumento, haciendo de su recolección, preproceso y análisis un gran reto, para quienes buscamos en los datos información de utilidad con fines comerciales o de investigación.

2.4 Las Redes Sociales en números

Volviendo al Global Digital 2019 [2], podemos señalar que las redes sociales constituyen una fuente importante de datos en el contexto de Big Data. Hoy en día el 45% de la población mundial es considerado usuario activo de las redes sociales, lo que significa la impresionante cantidad de 3,48 mil millones de personas (288 millones de usuarios nuevos en los últimos 12 meses). Y un hecho que es aún más interesante saber, es que los usuarios activos de las redes sociales en entornos móviles ya están en el 42%. A pesar de estas cifras, el uso de las redes sociales aún está lejos de estar distribuido de manera uniforme en todo el mundo como observamos en la figura 5.

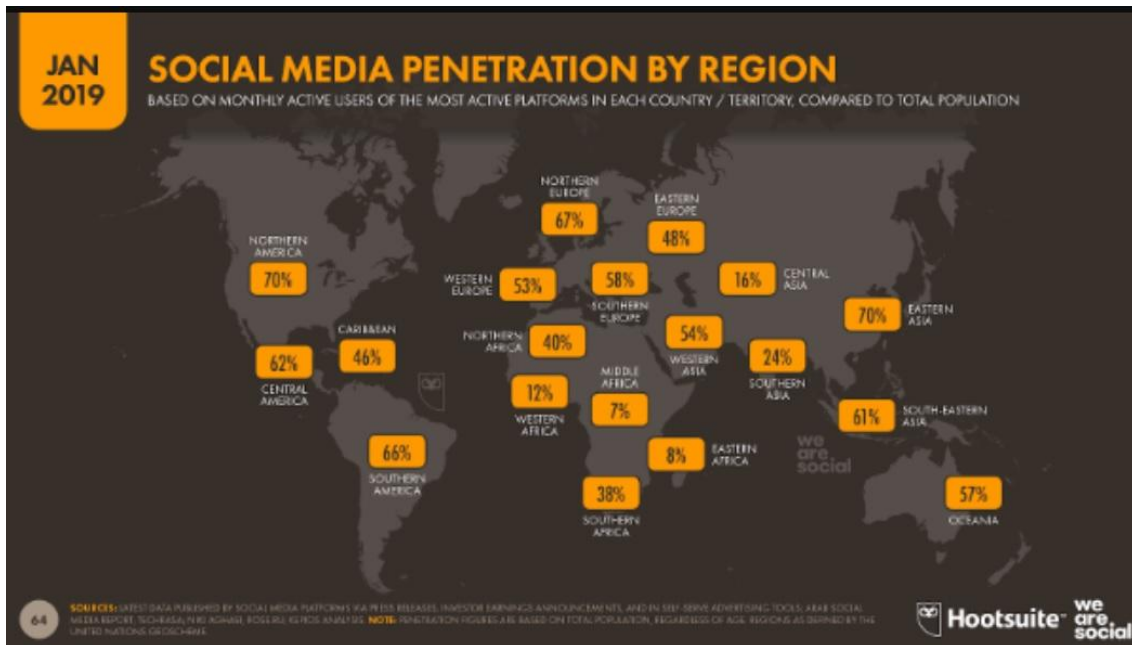


Figura 5. Penetración de las Redes Sociales por Región

En el Norte, Centro y Sudamérica, al igual que el Noreste europeo y en el Este/Sudeste asiático, se puede fácilmente visualizar que la penetración del uso de las redes sociales es mayor que en otras regiones.

Así mismo, la principal plataforma de redes sociales continúa siendo Facebook, con un incremento de usuarios activos en los últimos 12 meses de casi 10% de usuarios al mes (figura 6), mientras que Twitter, la red social que analizaremos en el presente trabajo de fin de Máster (TFM), a pesar que presenta una disminución del 1.5% de los usuarios activos desde octubre del 2018 [2], esta situación no ha mermado en la importancia que tiene dicha red social a nivel mundial (figura 7).

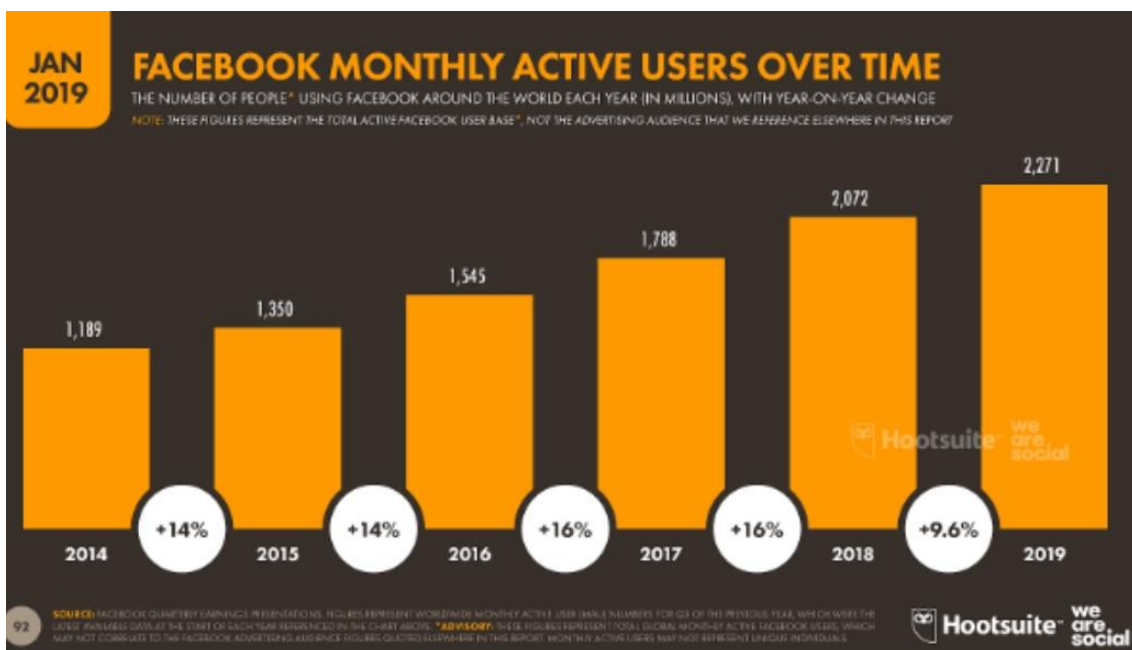


Figura 6. Usuarios activos mensualmente en Facebook

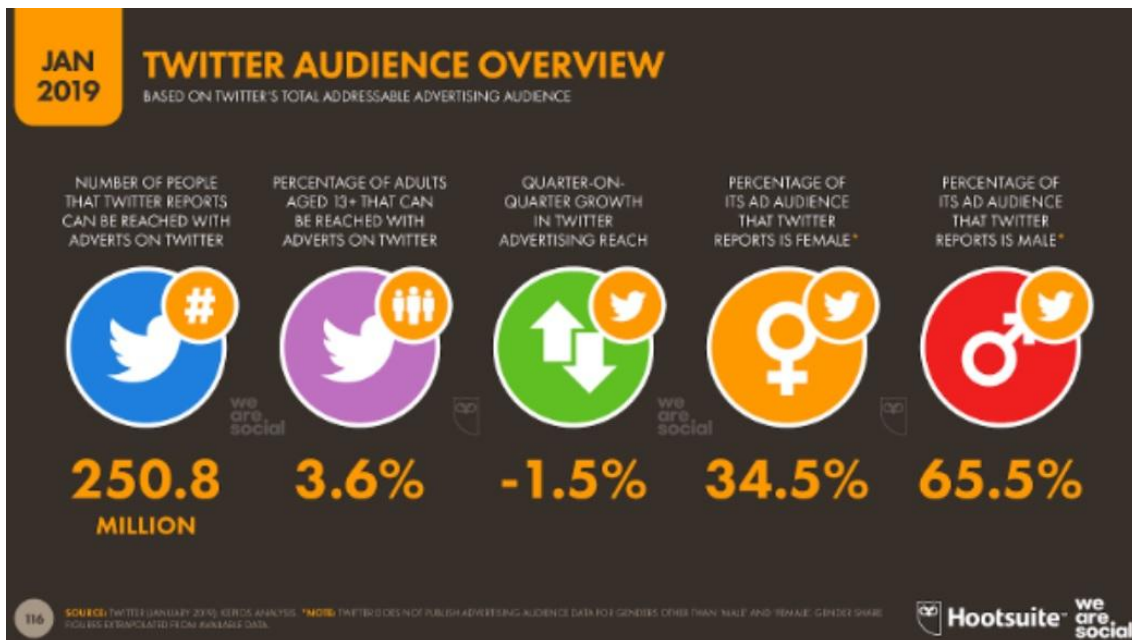


Figura 7. Visión general de la audiencia de Twitter

2.5 Twitter como Red Social

Por muchos, Twitter es considerado como "microblogging" o "nanoblogging" debido a que constituye un servicio que permite a sus usuarios enviar y publicar mensajes breves llamados tweets (hasta 280 caracteres), donde se puede incluir: texto, enlaces a sitios web, emoticones y otros recursos relevantes. Algunos no la consideran una red social, pero en este trabajo lo es, y es considerada una "red" porque interconecta individuos y organizaciones con intereses comunes dentro de la sociedad.

Los usuarios de Twitter "siguen" a otros usuarios. Si sigues a alguien, puedes ver sus tweets en la "línea de tiempo" de tu cuenta en Twitter. Puedes elegir seguir a personas y organizaciones con intereses académicos y personales similares a los tuyos. [11]

Puedes crear tus propios tweets o puedes retuitear información que otros han twitteado. Retuitear significa que la información se puede compartir de manera rápida y eficiente con un gran número de personas. [11]

En el proceso de conectar a personas que son completamente extrañas, en Twitter se puede utilizar hashtags. Los hashtags, que se indican con el prefijo "#", se agregan a los tweets para que los miembros de la comunidad puedan compartir la conversación. Los usuarios pueden incluso usar un sitio web como hashtag.org para encontrar temas que les interesen. Luego, pueden usar esos hashtags para unirse a las conversaciones que tienen lugar sobre el tema, y finalmente ayudar a construir comunidades en línea basadas en el contenido. [12]

En el presente trabajo se eligió la red social Twitter y no otra, básicamente por las siguientes razones:

1. Por la atención mediática que recibe, al ser una plataforma popular, los principales incidentes, noticias y eventos se comentan allí y tienden a estar centrados en torno a uno o más hashtags, haciendo fácil encontrar y seguir conversaciones a partir de éstos o también a partir de palabras claves.
2. Porque la recolección, clasificación, y la ampliación de la búsqueda de los tweets se realiza a través de la API de Twitter, que es más abierta y accesible en comparación con las de otras plataformas de medios sociales.
3. Porque muchos investigadores están utilizando Twitter como fuente primaria de recolección de datos, de esta forma, cada vez se encuentra mayor cantidad de estudios de investigación social basados en el contenido de esta red social.

Actualmente, Twitter posee 250 millones de usuarios a nivel mundial, de los cuales las tres cuartas partes pertenecen al sexo masculino. [1]

2.6 Twitter: Data Mining y Análisis de Redes

Twitter proporciona a ciertas empresas, desarrolladores y usuarios acceso a datos públicos a través de sus interfaces de programación de aplicaciones (API) o software que solicita y entrega información. La compañía, normalmente, vende los datos a organizaciones, que a menudo los usan para analizar eventos o servicio al cliente. [13]

En estos casos, los clientes empresariales reciben el acceso de datos más amplio, que incluye los últimos 30 días de tweets o acceso a tweets desde el año 2006. Para obtener ese acceso, los clientes deben explicar cómo planean usar los datos y quiénes son los usuarios finales. Además, Twitter no vende datos privados de mensajes directos. [13]

Para el desarrollo del presente trabajo, hemos accedido a los datos públicos de Twitter comprendidos entre el 4 al 12 de Marzo del 2019, utilizando los siguientes hashtags y palabras clave:

#8M
#Marcha8M
#DiaInternacionalDeLaMujer
#DiaDeLaMujer
#8marzo2019
#8M2019
Feminicidio
Acoso
Feminismo
Feministas

2.6.1 Terminología en Twitter

A continuación, el significado de los términos más utilizados en Twitter [20]:

- **Tweet** (Tuit): Publicación/mensaje o actualización de estado en Twitter.
- **Followers** (Seguidores): Usuarios que siguen una cuenta y leen los tweets que se envían.
- **Following** (Seguidos): Son las cuentas de Twitter que un usuario sigue. Los mensajes publicados en las cuentas seguidas aparecen automáticamente en el timeline del usuario.
- **Timeline** (TL): Lista de tweets de las cuentas que se siguen de manera cronológica.
- **@Usuario**: Es la forma en que se muestra el nombre del usuario en los mensajes que ha enviado. El preceder con una @ el nombre del usuario, es un enlace directo a su perfil de Twitter.
- **Listas**: Son agrupaciones de cuentas, seleccionadas generalmente por temática, que crea el usuario.
- **Perfil**: Es la página de Twitter de cada usuario donde se puede consultar la información del mismo, y los tuits que ha enviado.
- **Retweet** (RT o retuit): Retuitear es una función que permite publicar en tu cuenta de Twitter un tweet de otra persona, compartiéndolo con tus seguidores, porque por algún motivo te parece interesante o importante.
- **DM** (Mensaje Directo): Son mensajes privados entre 2 usuarios, no son públicos por lo cual no aparecen en la secuencia cronológica de eventos (timeline).
- **HashTag** (#): Son las palabras clave o etiquetas precedidas por el símbolo #. Esta funcionalidad permite unir conversaciones, ya que los demás usuarios pueden utilizarlo o facilitar la búsqueda de un tema.
- **Trending Topics** (Temas en boga): Son las palabras clave que se utilizan en un transcurso de tiempo y permiten ver que tan popular se volvió un tema. Normalmente su popularidad se reduce a horas o minutos.
- **#FF** (Follow Friday): Es un hashtag muy utilizado en Twitter los días viernes, para recomendar usuarios. Se escribe el hashtag seguido de las cuentas que invitas a seguir.
- **Avatar**: Es la imagen del perfil que representa a cada usuario y que generalmente acompaña los mensajes enviados por el mismo.

- **Menciones:** Son tuits en los que se menciona el nombre de un usuario y permite saber quien fue el usuario autor. La diferencia entre Mención y Reply es que, el segundo comienza con @NombredelUsuario, y en la mención no es necesario que el mensaje comience de esa manera. Es decir, todos los replies son menciones pero no todas las menciones son replies.
- **Sidebar:** Barra que contiene los botones que ayudan a navegar en un perfil de Twitter y a su vez muestran la información del usuario y su cuenta.
- **Unfollower:** Dejar de seguir a un usuario, como cancelar la suscripción a sus tuits, por lo cual éstos dejan de aparecer en el timeline.
- **Favoritos:** Tuits que cada usuario considera de interés. Se marcan con la estrella que aparece en la parte superior de cada tuit.
- **Bot:** Perfil robot que manda tuits programados.
- **Troll:** Usuario de Twitter que utiliza su cuenta para atacar a otros usuarios con comentarios agresivos o groseros.

2.6.2 Datos que se pueden extraer desde Twitter

Los datos más relevantes a extraer de Twitter son:

- **Quién.** La persona que escribe el tweet, o hizo un retweet, junto con sus datos públicos como: nombre completo, localización, lenguaje, etc.
- **Cuándo.** Fecha y hora de publicación. Por ejemplo, con el perfil del usuario se puede determinar el horario en el que se encuentra.
- **Qué.** El contenido o mensaje del tweet, que además de texto, incluye links, menciones o contenido multimedia.
- **Dónde.** Coordenadas geográficas de la ubicación desde dónde fue publicado el tweet. Este dato es opcional, por tanto, no todos los tweets contienen esta información.

Para la extracción de los datos desde Twitter, se ha utilizado la librería open source de Python llamada Tweepy, la cual provee acceso a los métodos de la API de Twitter (explicada en detalle en la documentación oficial ubicada en <https://dev.twitter.com/>), y los resultados obtenidos fueron almacenados en archivos de tipo JSON. Luego, estos archivos se leyeron y colocaron en una estructura de datos (Diccionarios o Pandas DataFrame), para proceder con el resto del proceso de pre-procesado de datos (eliminación de puntuaciones, acentos, espacios en blanco, unicodes, URL, #nombre de usuarios, emojis, RT, etc.).

2.6.3 Redes de Retweets y métricas

En cuanto al análisis de redes, se ha considerado crear redes basadas en Retweets (RT), utilizando técnicas para visualizar los enlaces entre las personas, para comprender mejor la estructura de las conversaciones y para darnos una gran perspectiva sobre las personas en la red (quienes son personas influyentes reales, quienes están más conectados, etc.).

Así, para el diseño de cada red, los datos a considerar son:

Origen: Usuario que escribe el tweet.

Destino: Usuario/s que retuitea el tweet del usuario origen.

En la figura 8, los nodos o círculos serán los usuarios y las líneas que unen dichos nodos serán los retweets. De este modo, cada usuario se considera un nodo, y cada retweet un enlace entre estos nodos. Por ejemplo, si el usuario B retuitea al usuario A en un tweet, se establece que existe una relación entre A y B.

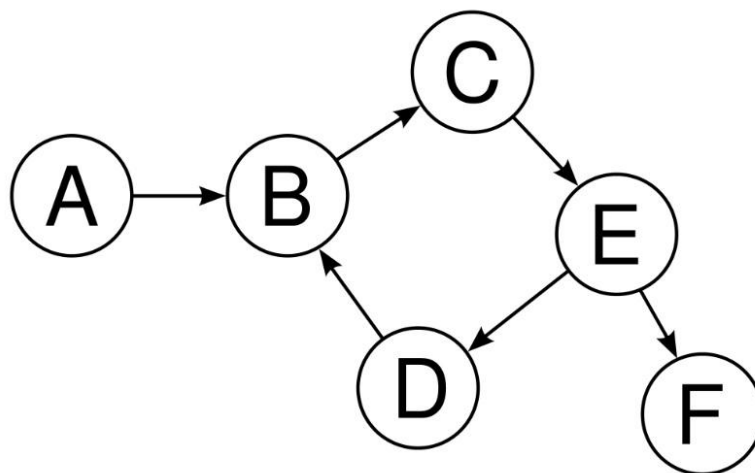


Figura 8. Gráfico de red dirigida básica

Los enlaces entre los usuarios de ejemplo A, B, C, D, E y F en la figura 8, configuran el grafo general, y sobre este grafo es donde realizaremos el análisis.

Concretamente: los gráficos son estructuras matemáticas que se utilizan para estudiar relaciones de pares entre objetos y entidades. Es una rama de las Matemáticas Discretas y ha encontrado múltiples aplicaciones en Informática, Química, Lingüística, Investigación de Operaciones, Sociología, etc. [14]

Adicionalmente, las redes se clasifican como dirigidas o no dirigidas, una distinción que se relaciona con el concepto de nodos de origen y destino. En una red dirigida, los enlaces representan una interacción unidireccional de un nodo a otro, mientras que en una red no dirigida, no hay distinción entre los nodos conectados (los enlaces son mutuos y no tienen "dirección" como tal).

La red se puede modelar usando diferentes herramientas, por ejemplo en el lenguaje "R" y/o Python para analizar datos usando conceptos de la teoría de grafos. Para este trabajo se utilizaron las librerías de Python: Ujson y Networkx. Ujson para la carga rápida de los archivos de datos de tipo Json, y Networkx porque permite las importaciones directas de Pandas Dataframe. Adicional, y paralelamente, se utiliza Gephi como herramienta de código abierto (open source) para la visualización y análisis de redes.

Las herramientas de análisis utilizadas calcularon las siguientes métricas:

2.6.3.1 La Centralidad

Esta es una de las herramientas conceptuales más utilizadas e importantes para el análisis de redes. La centralidad o Centrality tiene como objetivo encontrar los nodos (o usuarios) más importantes en una red. Las medidas de centralidad pueden ayudarnos a identificar la popularidad, los gustos más populares y las personas más influyentes dentro de la red. [14] Puede haber diferentes nociones de "importante" y, por lo tanto, hay muchas medidas de centralidad. Las medidas de centralidad en sí mismas tienen una forma de clasificación (o tipos de medidas de centralidad). Hay medidas que se caracterizan por el flujo a lo largo de líneas o enlaces y aquellas que se caracterizan por la estructura del camino.

Algunos de los tipos de centralidad más utilizados son:

1. **Centralidad de Grado (Degree Centrality):** Este es el número de enlaces conectados a un nodo. En el caso de una gráfica dirigida, podemos tener medidas de centralidad de 2 grados. Centralidad de entrada y salida.
2. **Centralidad de proximidad (Closeness Centrality):** es la longitud promedio de la ruta más corta desde un nodo hacia todos los demás nodos.
3. **Centralidad de intermediación (Betweenness Centrality):** la cantidad de veces que un nodo está presente en la ruta más corta entre otros 2 nodos.
4. **Centralidad de vector propio (Eigenvector Centrality):** mide cómo está conectado un nodo y la posible influencia directa que puede tener en la mayoría de nodos activos en la red.

Estas medidas de centralidad tienen variantes y las definiciones se pueden implementar utilizando varios algoritmos. En definitiva, esto significa un gran número de definiciones y algoritmos.

2.6.3.2 La Densidad de la red

Una medida de cuántas líneas o enlaces tiene un gráfico. La definición real variará según el tipo de gráfico y el contexto en el que se formule la pregunta. Para un gráfico completo no dirigido, la densidad es 1, mientras que es 0 para un gráfico vacío. La densidad de gráfico puede ser mayor que 1 en algunas situaciones (que involucran bucles). Una colección más densa de nodos implica una mayor interconectividad. La escasez implica una interconectividad más débil.

2.6.3.3 Coeficiente de agrupamiento

Se observa que las personas que comparten conexiones en una red social tienden a formar asociaciones. En otras palabras, hay una tendencia en una red social a formar grupos (clusters). Podemos determinar los grupos de un nodo, el coeficiente de agrupación local, que es la fracción de los pares de amigos del nodo (es decir, las conexiones) que están conectados entre sí. [15] A partir del agrupamiento es posible la detección de comunidades y entre los más innovadores métodos de detección de comunidades, Girvan y Newman[16] tiene un papel principal.

Dentro de la red, hemos mencionado que es posible determinar qué perfiles de usuarios (personas o nodos) son los más relevantes e influyentes, y que también es posible identificar a las comunidades (clusters) que conforman dicha red, para ello, en el presente trabajo, utilizando la herramienta Gephi y el lenguaje de programación Python, se realizarán las siguientes tareas:

a) Distribuir los nodos

En Gephi, se utilizarán dos algoritmos dirigidos por la fuerza:

- **Force Atlas 2.** Que simula un sistema físico para espacializar una red. Los nodos se rechazan entre sí como partículas cargadas, mientras que los enlaces atraen a sus nodos, como los resortes. Estas fuerzas crean un movimiento que converge a un estado equilibrado.
- **Fruchterman-Reingold.** Que simula la gráfica como un sistema de partículas de masa. Los nodos son las partículas de masa y los enlaces son resortes entre las partículas que se atraen como imán. Si bien en algunas gráficas su ejecución ha resultado un poco lenta, el resultado final obtenido bien justifica el tiempo de espera.

b) Determinar los nodos influyentes

En Python se calcularán las métricas de centralidad de Grado (Degree), de Proximidad (Closeness Centrality) y de Intermediación (Betweenness Centrality). En Gephi obtendremos la centralidad de vector propio (Eigenvector Centrality), algoritmo que utilizaremos para identificar a los nodos influyentes.

c) Identificar las comunidades

Se aplicará el algoritmo de clustering llamado “Modularity” incluido en Gephi, y con él se identificarán los grupos de usuarios que por algún motivo se asemejen y mantengan más enlaces entre sí que con el resto de nodos de la red.

2.7 Twitter como recurso feminista

Desde la aparición del feminismo a finales del siglo XIX, este ha cobrado importancia desde mediados del siglo XX, como un movimiento político, social e ideológico que plantea que la mujer debe tener en la sociedad actual el derecho a:

- Cumplir el mismo rol que el hombre.
- Participar políticamente.
- Trabajar, a estudiar, a decidir sobre su cuerpo y su vida.
- Construir su identidad de manera autónoma y sin imposiciones ni de otros, sean hombres, instituciones formales, la Iglesia o el Estado.

En muchos países se vienen gestando innumerables actos (eventos, marchas, conferencias, protestas, etc.), con el fin de concientizar a la población sobre la existencia de un sinnúmero de situaciones en las cuales las mujeres sufren de algún tipo de violencia, ejercicio del poder o dominio sobre sus cuerpos, mentes y elecciones, ya sea tanto a manos de hombres o incluso a manos de otras mujeres, lo cual habla de que aún las sociedades occidentales pueden tener todavía rastros de machismo ocultos o no tan evidentes que dañan y perjudican a la construcción de la identidad de cada una de las mujeres. [17]

En la actualidad, las redes sociales constituyen una herramienta importante que reúne a las comunidades de toda índole de intereses y el movimiento feminista no escapa a ello. A través de las redes sociales podemos enviar datos de un extremo del mundo a otro en cuestión de segundos, hacer presentaciones en línea, usar imágenes, videos, sonidos y textos para compartir nuestras vidas reales: hoy las historias personales se hacen públicas y los problemas locales se vuelven globales. Y un problema local en Perú (pero global en todo el mundo), lo constituye la permanente lucha por los derechos de la mujer.

En los Estados Unidos, a partir del estudio de Jackson y Banaszczyk [18] del 2016, se examina el trabajo discursivo y los debates conformados por los hashtags #YesAllWomen y #YesAllWhiteWomen, identificando a los usuarios más influyentes de estos hashtags en las redes y analizando críticamente sus tweets. Constatan también como los grupos feministas utilizan Twitter para producir y comunicar temas sobre la violencia contra las mujeres y al mismo tiempo participar en debates comunitarios sobre raza e inclusión. Por otro lado, Barker-Plummer [19] en el 2017 analiza también el hashtag #YesAllWomen, como uno de los eventos de feministas más grandes y más visibles de los últimos años en Twitter, y lo considera un ejemplo importante del activismo discursivo feminista contemporáneo, concluyendo que dicho hashtag funcionó,

primero, como un sitio de identidad colectiva para los participantes, y segundo, que a través de sus enlaces y recirculación por otras plataformas y los medios de comunicación, #YesAllWomen también funcionó como una protesta pública o un evento de creación de agenda con impacto en el discurso público más allá de Twitter.

En Sudamérica, Alcaraz [21] analiza el movimiento #NiUnaMenos en Twitter, que comenzó en Argentina como reacción al asesinato en 2015 de una niña argentina de 14 años por parte de su novio y que en la actualidad se ha extendido por el resto de América Latina. El hashtag #NiUnaMenos, ampliamente conocido en Perú, significa "**no más femicidios**", pero también es utilizado como una demanda para el fin de todas las formas de opresión, desde la extrema violencia y abortos inseguros, para trabajos de cuidado no remunerados y hostigamiento en la calle. La autora, considera que existe "una alianza entre tecnología, redes sociales y personas en las calles". [20]

3. Minería de Datos

Durante el desarrollo del presente trabajo, la etapa de Minería de Datos, es la que mayor tiempo de dedicación ha demandado (alrededor de 21 días), comprendiendo las siguientes tareas:

3.1 Programación en Python

3.1.1 Programa para captura de tweets por fecha y palabra clave

Se empleó una cuenta de desarrollador previamente registrada en Twitter así como una biblioteca de código abierto de Python llamada Tweepy, a fin de acceder mediante las claves de acceso y tokens a las API de Twitter desde un programa desarrollado en Python. Como resultado de su ejecución, se obtuvieron los tweets durante el período del 4 al 12 de marzo del 2019. Dichos datos, correspondientes a tweets recogidos por fecha y palabra clave se almacenaron en archivos de tipo Json.

Durante el proceso de captura se señaló como geo-localización inicial la de Perú: latitud -9.189967, y longitud -75.015152, considerando también un espacio de 2500 km alrededor de dicho punto. Los archivos de tipo Json obtenidos tuvieron un tamaño promedio de 1GB cada uno y su estructura se puede visualizar en la figura 9.



Figura 9. Estructura de archivo de Tweets (de tipo Json)

3.1.2 Programa para lectura y proceso de archivos de tweets

La relación con respecto al total de tweets capturados versus aquellos que pertenecen o se encuentran registrados en territorio peruano fue de 1:150, es decir, que por cada 150 tweets recogidos, solo un tweet correspondía como enviado desde Perú.

En Python, los tweets contenidos en archivos de tipo Json fueron leídos mucho más rápido a través de la biblioteca Ujson (en lugar de la librería Json) y se cargaron a una estructura denominada dataframe perteneciente a la biblioteca Pandas. Bajo dicha estructura de datos se aplicaron filtros y se eliminaron datos para finalmente generar archivos de tipo CSV, conteniendo sólo los datos de los tweets requeridos para el modelado de las redes.

3.1.3 Programa para creación de archivos para grafos

Como veremos en el capítulo 4 de modelado de red, la herramienta Gephi que utilizaremos, deberá importar una 'Tabla de nodos' y una 'Tabla de enlaces' por separado (en formato CSV), por ello, la necesidad de crear estos dos archivos:

- Archivo de Nodos
- Archivo de Interacciones (enlaces)

3.1.3.1 Archivo de Nodos

Los nodos están representados por las personas o actores que escriben los tweets, así como sus principales características. En este archivo se incluyeron los siguientes datos:

- ID de usuario de Twitter
- Atributos
 - Etiqueta (nombre de pantalla)
 - Fecha de creación del usuario
 - Imagen de perfil
 - Número de seguidores
 - Número de amigos

Algunos nodos obtenidos se observan en la figura 10.

	A	B	C	D	E	F
1	Id	Label	user_created_at	profile_image	followers_count	friends_count
2	1.0033E+18	saladademrd	Sun Jun 03 15:56:54 +0000 2018	https://pbs.twimg.com/profile_images/11034592467237	196	243
3	100373200	estefanyfrancoa	Wed Dec 30 01:52:14 +0000 2009	https://pbs.twimg.com/profile_images/10451816732420	360	516
4	1.0042E+18	BrunelloInfante	Wed Jun 06 03:59:31 +0000 2018	https://pbs.twimg.com/profile_images/10607115441898	8	207
5	1.0043E+18	Analucia_1410	Wed Jun 06 12:52:19 +0000 2018	https://pbs.twimg.com/profile_images/11058150465784	52	106
6	100561404	postmoderngrll	Wed Dec 30 18:17:50 +0000 2009	https://pbs.twimg.com/profile_images/95525708733177	277	140
7	1.0063E+18	AngelLe27531858	Tue Jun 12 01:20:34 +0000 2018	https://pbs.twimg.com/profile_images/10734511856994	26	182
8	1.0069E+18	lexithalia	Wed Jun 13 16:03:29 +0000 2018	https://pbs.twimg.com/profile_images/10848562478353	66	266
9	1.0077E+18	uchi_fu1	Fri Jun 15 16:59:16 +0000 2018	https://pbs.twimg.com/profile_images/10248564105286	26	652
10	1.0078E+18	sunflowercly	Fri Jun 15 23:22:43 +0000 2018	https://pbs.twimg.com/profile_images/11045397309878	35	142
11	100782421	jesusnoriega	Thu Dec 31 14:58:47 +0000 2009	https://pbs.twimg.com/profile_images/10783292580139	1384	2256
12	1.0082E+18	vaalcalp	Sun Jun 17 07:07:36 +0000 2018	https://pbs.twimg.com/profile_images/10082483745990	23	48
13	1.0084E+18	pescadx_rabioso	Sun Jun 17 17:36:44 +0000 2018	https://pbs.twimg.com/profile_images/10930342422238	284	106
14	1.0087E+18	claudiacrmn	Mon Jun 18 15:50:40 +0000 2018	https://pbs.twimg.com/profile_images/10979631275121	85	512

Figura 10. Estructura de datos de archivo de nodos (de tipo Csv)

3.1.3.2 Archivo de Interacciones (enlaces)

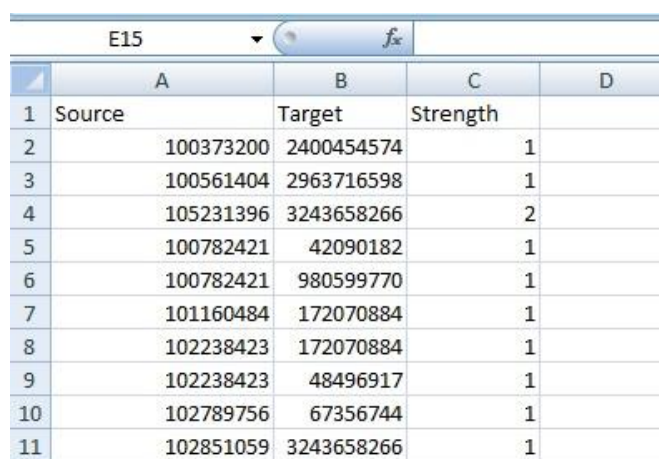
Este archivo contiene el nodo origen y el nodo destino, como así lo sugieren las palabras 'fuente' y 'objetivo', reflejando de esta forma la relación entre pares de nodos. Adicionalmente, se ha contemplado el atributo "Strength" o fuerza, traducida en la cantidad de veces en total que cada uno de los usuarios retweetó al otro (no se han considerado las menciones ni las respuestas a otros usuarios). Entonces, en este archivo se incluyen los siguientes datos:

- Source o Nodos de origen (ID de usuario de Twitter)
- Target o Nodos de destino (ID de usuario de Twitter)
- Atributo
 - Strength o Fuerza

Para el presente trabajo, se tomó como atributo de fuerza (Strength) el valor de "1", esto quiere decir, que bastó que un tweet tuviera por lo menos un ReTweet para ser considerado.

Por otro lado, si se hubiera tomado como atributo de fuerza por ejemplo, el valor de "5" como mínimo, sólo se hubieran filtrado aquellos tweets que tuvieran cinco ReTweets o más, seleccionando entonces los tweets con las relaciones más fuertes (y por consiguiente, un resultado menor en cuanto a la cantidad de nodos que cumplieran con dicha condición).

Algunas interacciones obtenidas se observan en la figura 11.



	A	B	C	D
1	Source	Target	Strength	
2	100373200	2400454574	1	
3	100561404	2963716598	1	
4	105231396	3243658266	2	
5	100782421	42090182	1	
6	100782421	980599770	1	
7	101160484	172070884	1	
8	102238423	172070884	1	
9	102238423	48496917	1	
10	102789756	67356744	1	
11	102851059	3243658266	1	

Figura 11. Estructura de datos de archivo de interacciones (de tipo Csv)

La relación de archivos de tipo CSV, correspondientes a nodos e interacciones generados, se muestra en la figura 12.

Name	Date modified	Type	Size
Videos	02-05-19 9:28 PM	File folder	
edges_#8M_pe_2019-03-04_to_2019-03-07	02-05-19 9:36 PM	Archivo de valores...	36 KB
edges_#8M_pe_2019-03-08	02-05-19 7:16 PM	Archivo de valores...	39 KB
edges_#8M_pe_2019-03-09_to_2019-03-12	02-05-19 11:02 PM	Archivo de valores...	32 KB
edges_#acoso_04_to_12_mar_2019	03-05-19 7:24 PM	Archivo de valores...	284 KB
edges_#feminicidio_04_to_12_mar_2019	03-05-19 8:40 PM	Archivo de valores...	15 KB
edges_#feminismo_04_to_12_mar_2019	03-05-19 8:06 PM	Archivo de valores...	52 KB
edges_#feminismo_y_feministas_04_to_12_mar_2019	03-05-19 7:46 PM	Archivo de valores...	93 KB
edges_#feministas_04_to_12_mar_2019	03-05-19 8:19 PM	Archivo de valores...	48 KB
edges_#Marcha8M_04_to_12_mar_2019	03-05-19 11:00 PM	Archivo de valores...	196 KB
nodes_#8M_pe_2019-03-04_to_2019-03-07	02-05-19 9:36 PM	Archivo de valores...	273 KB
nodes_#8M_pe_2019-03-08	02-05-19 7:16 PM	Archivo de valores...	364 KB
nodes_#8M_pe_2019-03-09_to_2019-03-12	02-05-19 11:02 PM	Archivo de valores...	330 KB
nodes_#acoso_04_to_12_mar_2019	03-05-19 7:24 PM	Archivo de valores...	1,236 KB
nodes_#feminicidio_04_to_12_mar_2019	03-05-19 8:40 PM	Archivo de valores...	167 KB
nodes_#feminismo_04_to_12_mar_2019	03-05-19 8:06 PM	Archivo de valores...	431 KB
nodes_#feminismo_y_feministas_04_to_12_mar_2019	03-05-19 7:46 PM	Archivo de valores...	655 KB
nodes_#feministas_04_to_12_mar_2019	03-05-19 8:19 PM	Archivo de valores...	377 KB
nodes_#Marcha8M_04_to_12_mar_2019	03-05-19 11:00 PM	Archivo de valores...	1,462 KB

Figura 12. Relación de archivos de nodos y enlaces (de tipo Csv)

3.2 Filtrado y eliminación de datos

Como cada archivo de tipo Json contenía los tweets no sólo de Perú sino también los tweets generados a una distancia de 2500km a la redonda (abarcando parte de Ecuador, Chile, Argentina y Brasil), estos archivos en su mayoría alcanzaron o sobrepasaron 1GB de tamaño, por ello, se tuvieron que filtrar y eliminar los datos innecesarios (caso contrario, nuestros modelos de redes hubieran contemplado muchos más nodos y enlaces, no necesariamente pertenecientes a Perú).

El filtrado y eliminación de datos se llevó a cabo en base al siguiente criterio:

- Se seleccionó como válidos sólo aquellos tweets cuya localización (longitud/latitud) figuraban dentro de Perú.
- Se revisó el campo "location", seleccionando sólo aquellos tweets en donde dicho campo de texto contenía la palabra Perú.
- Como nuestras redes están basadas en los Retweets, se seleccionaron sólo aquellos tweets que por lo menos hicieron "1" Retweet (RT), a fin de obtener la mayor cantidad de nodos y relaciones. Si en vez de "1", se hubiera contemplado una fuerza de "5" Retweets como mínimo, en este caso, los nodos y sus relaciones hubieran disminuido drásticamente en número para el modelado de las redes, porque sólo se hubieran mostrado los nodos y relaciones más fuertes.

4. Modelado y análisis de las Redes

Bajo el modelo de análisis de redes sociales (SNA), como ya mencionáramos anteriormente, las redes sociales están representadas por "nodos" (por ejemplo, individuos, organizaciones, ideas) y "enlaces" o las diversas relaciones que los conectan. En el caso de las redes sociales en Twitter, los nodos pueden representar diferentes usuarios, mientras que los enlaces pueden representar cualquiera de las interacciones que conectan estas cuentas individuales (por ejemplo, me gusta, respuestas, retweets, menciones, etc.).

Teniendo en cuenta el SNA, para el modelado y visualización de nuestras redes, se ha utilizado el software llamado Gephi en su versión 0.9.2, que es una plataforma interactiva de código abierto y está disponible para descarga gratuita en línea, mientras que para la ejecución de las pruebas estadísticas esenciales se empleó el lenguaje de programación Python 3.7.

La elección de Gephi para la visualización, se justifica porque es una herramienta muy amigable y viene siendo utilizada con frecuencia en los análisis de redes relacionados con las ciencias sociales, estudios de índole cultural, etc., y por otro lado, se eligió el lenguaje Python, por su robustez, potencia y sencillez para el cálculo estadístico.

En el presente trabajo, es importante señalar también que:

- Las redes se modelan a partir de los Retweets como tipo de interacción entre nodos.
- Todas las redes serán de tipo dirigidas, es decir, que sus enlaces representan una interacción unidireccional de un nodo a otro.
- Las métricas a calcular son las siguientes:
 - **Densidad de la red.** Es una medida que toma valores entre cero y uno e indica la proporción de vínculos presentes entre los actores, frente al número total de conexiones posibles.
 - **Centralidad.** Se ha identificado a los diez nodos más influyentes a partir del cálculo de las siguientes métricas de centralidad:
 - Centralidad de grado o Degree
 - Centralidad de intermediación o Betweenness
 - Centralidad del vector propio o Eigenvector

La centralidad del vector propio o Eigenvector fue la métrica utilizada para la visualización de los nodos en cada una de las redes.

Se justifica el cálculo adicional de las otras métricas de centralidad (la de grado y de intermediación) por las siguientes razones:

- Se constata que la identificación de los nodos influyentes según estos tipos de centralidad no necesariamente coinciden con los obtenidos según la centralidad de vector propio. Mientras la centralidad de grado seleccionó aquellos nodos con la mayor cantidad de relaciones/enlaces dentro de la red, la centralidad de intermediación hizo la selección de los nodos influyentes, según la importancia de un nodo para las rutas más cortas a través de la red.
- Se calcula la centralidad de grado sin tener en cuenta la estructura global de la red.
- Al calcular la centralidad de intermediación, se identifican actores que hacen de puente entre otras regiones de la red, por ello, pueden conectar distintas comunidades entre sí. De esta forma, si conseguimos que uno de estos actores influyentes nos mencione o nos haga RT, nuestro tweet podría llegar a otro tipo de usuarios que quizás estén interesados en nuestros temas.
- **Modularidad de la red.** A partir de la función Modularidad dentro de la plataforma Gephi se identifican a las Comunidades dentro de cada red.
- Las métricas que no fueron contempladas en nuestro análisis son:
 - El mayor sub-grafo conexo de la red.
 - El camino más corto (geodésico) en el mayor subgrafo conexo de la red y su longitud.
 - Diámetro, radio y media del componente más grande de la Red.
- La visualización de las redes considera lo siguiente:
 - El tamaño de los nodos refleja cuán importantes son para la red. Se ha utilizado la centralidad Eigenvector o de vector propio para identificar a los nodos más influyentes y visualizarlos en cada una de las redes. Al distinguir la importancia de los nodos según la **centralidad de vector propio** (centralidad Eigenvector), se obtuvo un valor de 0 a 1: cuanto más cerca de uno, mayor es la centralidad. Este tipo de centralidad es de utilidad para conocer qué nodos pueden obtener información de muchos otros nodos de forma rápida (algo parecido a la extensión utilizada por Google en su algoritmo de PageRank para determinar qué páginas web llegan a los primeros resultados de búsqueda).

La selección de la centralidad de vector propio (Eigenvector) para la visualización de los nodos influyentes se justifica por las siguientes razones:

- Porque mide la importancia de un nodo al tiempo que considera la importancia de sus vecinos.

- Toma en consideración no sólo cuántas conexiones tiene un vértice (es decir, su grado), sino también el grado de los vértices a los que está conectado.
- Porque desarrolla una visión más sofisticada de la centralidad: una persona con pocas conexiones podría tener una centralidad muy alta del vector propio si esas pocas conexiones estuvieran muy bien conectadas.
- Porque es una medida muy similar a la que usa Google para clasificar las páginas web por importancia (el algoritmo de PageRank utilizado por el motor de búsqueda de Google).
- Modularidad de la red. Se han agrupado los nodos por colores que pertenecen a una misma Comunidad. La modularidad mide la cantidad de enlaces dentro de las comunidades, en comparación con la cantidad de enlaces entre comunidades. Los nodos se han coloreado en función de sus agrupaciones (a la comunidad que pertenece), según el resultado de la modularidad como parámetro de la partición
- Visualización de dos tipos de diseño de grafos:
 - **Diseño Fruchterman-Reingold**

Este es un algoritmo de gráfico dirigido por la fuerza y nos ofrece una visión clara de las conexiones en nuestra red. Adicionalmente, nos ayudará en el siguiente apartado a fin de facilitar la visualización de las Comunidades.
 - **Diseño Force Atlas 2**

Este es un algoritmo de gráfico dirigido que asume que los bordes son resortes y los nodos son imanes que se repelen entre sí. Entonces, los nodos estarán bien dispersos pero los enlaces no serán demasiado largos. Así, al simular este sistema de resorte / imán, se obtiene una buena imagen del grafo.
- La identificación de las Comunidades en cada red considera lo siguiente:
 - Se describen algunas observaciones de las principales comunidades detectadas en cada red a partir de la función de "Modularidad" dentro la plataforma Gephi.
 - Se detalla a nivel de las relaciones dentro de las comunidades, con el objetivo de determinar cuán "activo" y qué tan conectado estuvo un usuario, considerado influyente, con otros.
 - Gephi asigna el mismo color a los nodos que comparten alguna relación de afinidad (a los de la misma comunidad).

- Las principales Comunidades detectadas son las siguientes:
 - ❖ Política
 - ❖ Periodismo
 - ❖ Gobierno
 - ❖ Organizaciones de Derechos Humanos y Feministas
 - ❖ Organizaciones Deportivas / Culturales
 - ❖ Otros

- Los resultados obtenidos al analizar las comunidades se lograron explorando los gráficos, haciendo mucho zoom, desplazamiento, lectura de etiquetas y búsqueda de más información de los nodos. Con ayuda del navegador se revisaron las cuentas de Twitter y se comprobó la pertenencia de dicho nodo a la comunidad asignada.

Teniendo delimitado los alcances de cálculo para nuestros modelos de redes e identificación de las comunidades, se procedió al análisis de las cinco redes propuestas, todas ellas circunscritas dentro del territorio peruano y que se listan a continuación:

- 4.1. Red 8 de Marzo
- 4.2. Red Día de la Mujer
- 4.3. Red Marcha 8M
- 4.4. Red de Acoso
- 4.5. Red de Femicidio

4.1 Red 8 de Marzo

Esta red corresponde al evento de conmemoración del día de la Mujer conocido con el Hashtag #8M

La red del 8 de Marzo se ha dividido en tres subredes o grafos, de acuerdo a la consideración de tres periodos de tiempo:

- 4.1.1. Grafo de la red 8M del 4-7 marzo
- 4.1.2. Grafo de la red 8M sólo el día central 8 marzo
- 4.1.3. Grafo de la red 8M del 9-12 marzo

4.1.1 Grafo de la red 8M del 4-7 marzo

A manera de explicación, en este primer grafo se detallará el proceso de obtención de las principales métricas en Python, mientras que en los siguientes grafos de red sólo se mostrará el resumen de los resultados obtenidos.

4.1.1.1. Principales métricas

Si bien Gephi nos proporciona opciones de cálculo de métricas para redes, se decidió calcularlas en Python, según el procedimiento que se detalla a continuación.

- Como primer paso, se importan las bibliotecas necesarias en Python, entre ellas la biblioteca Networkx:

```
In [1]: %matplotlib inline
import pandas as pd
import networkx as nx
import matplotlib.pyplot as plt
import numpy as np
import ast
import re

import sys
import warnings

if not sys.warnoptions:
    warnings.simplefilter("ignore")
```

- Se carga el grafo generado en Gephi (archivo 8M-04-al-07-Marzo-19.graphml):

```
In [2]: #Carga de archivo 8M del 4 al 7 de Marzo 2019
g = nx.read_graphml("8M-04-al-07-Marzo-19.graphml")
```

```
In [3]: print("1. El grafo 'g' tiene: " + str(len(g.nodes)) + ' nodos')
print("2. El grafo 'g' tiene: " + str(len(g.edges)) + ' aristas')
```

```
1. El grafo 'g' tiene: 1091 nodos
2. El grafo 'g' tiene: 1322 aristas
```

- Se convierte temporalmente el grafo a “no dirigido” a fin de poder calcular la densidad de la red:

```
In [5]: # Conversion a grafo no dirigido a fin de calcular la densidad de la red
g_undirected = g.to_undirected()
print("1. El grafo 'g_undirected' tiene: " + str(len(g_undirected.nodes)) + ' nodos')
print("2. El grafo 'g_undirected' tiene: " + str(len(g_undirected.edges)) + ' aristas')
density_g = nx.density(g_undirected)
print("3. Densidad de la Red:", density_g)
```

```
1. El grafo 'g_undirected' tiene: 1091 nodos
2. El grafo 'g_undirected' tiene: 1314 aristas
3. Densidad de la Red: 0.00220990758415392
```

- A continuación, se calcula tres de las medidas de centralidad más comunes:

(1) Grado o Degree

- Se calcula el grado de cada nodo de la red y se identifican aquellos 10 nodos/personas con la mayor cantidad de edges o conexiones.

```
In [14]: #Calculamos el grado al que llamaremos 'degree' y Los agregamos como un atributo de cada nodo de La red
degree_dict = dict(g.degree(g.nodes()))
nx.set_node_attributes(g, degree_dict, 'degree')
```

```
In [37]: #Ahora debemos clasificar y ordenar Los nodos por su grado para visualizar Los mas relevantes
#Utilizamos La funcion sorted() para ordenar Los diccionarios por sus claves y valores para asi
#encontrar Los 10 nodos principales clasificados por grado.
```

```
import operator

sorted_degree = sorted(degree_dict.items(), key=operator.itemgetter(1), reverse=True)

print("Los 10 nodos mas relevantes por Grado (personas con mas conexiones en la red social) son:")
for d in sorted_degree[:10]:
    print(g.node[d[0]]['label'], d[1])
```

```
Los 10 nodos mas relevantes por Grado (personas con mas conexiones en la red social) son:
EsMomento_pe 91
duhnia 88
lamula 70
MarisaGlave 56
emmaylola 55
paromujeresperu 45
NUMSTLosOlivos 40
amnistiaperu 37
CMPFloraTristan 37
Actua_Pe 30
```

(2) Centralidad Betweenness

```
In [17]: #Esta medida de centralidad no necesita recibir una Lista de nodos, solo el gráfico "g"

betweenness_dict = nx.betweenness_centrality(g) # Ejecutamos La centralidad betweenness

# Asignamos el valor obtenido en estas centralidades como un atributo a cada nodo de La red
nx.set_node_attributes(g, betweenness_dict, 'betweenness')
```

```
In [40]: #(2)Ordenamos Los valores de centralidad betweenness

sorted_betweenness = sorted(betweenness_dict.items(), key=operator.itemgetter(1), reverse=True)

print("Los 10 nodos/personas mas relevantes por Centralidad Betweenness son:")
for b in sorted_betweenness[:10]:
    print(g.node[b[0]]['label'], b[1])
```

```
Los 10 nodos/personas mas relevantes por Centralidad Betweenness son:
NUMSTLosOlivos 0.00129133317901736
paromujeresperu 0.0010298939726240256
Actua_Pe 0.0010020769806702529
CMPFloraTristan 0.0008714780621925468
paridadperu 0.00046225991597134204
DemusPeru 0.0003875731245761829
emmaylola 0.00022469622553025382
femimisha 0.00020509916673386235
NuevoPeruGP 0.00020372717203409444
promsex 0.00015414881970344854
```

(3) Centralidad Eigenvector

```
In [17]: #Esta medida de centralidad no necesita recibir una Lista de nodos, solo el gráfico "g"

eigenvector_dict = nx.eigenvector_centrality(g) # Ejecutamos La centralidad eigenvector

# Asignamos el valor obtenido en estas centralidades como un atributo a cada nodo de La red
nx.set_node_attributes(g, eigenvector_dict, 'eigenvector')
```

```
In [41]: #(3)Ordenamos Los valores de centralidad eigenvector
sorted_eigenvector = sorted(eigenvector_dict.items(), key=operator.itemgetter(1), reverse=True)

print("Los 10 nodos/personas mas relevantes por Centralidad Eigenvector son:")
for b in sorted_eigenvector[:10]:
    #print(b)
    print(g.node[b[0]]['label'], b[1])
```

Los 10 nodos/personas mas relevantes por Centralidad Eigenvector son:
 amnistiaperu 0.3730404935191632
 NUMSTLosOlivos 0.3354579611081344
 paromujeresperu 0.30094615430878097
 InfoLaPlaza 0.2780094797617802
 Actua_Pe 0.2626336185890785
 MadresAccionPe 0.2520057202822941
 emmaylola 0.22541018092755888
 NoticiasSER 0.1983619901328991
 Graciela_Jimena 0.1964647080124586
 CMPFloraTristan 0.18206943810530238

4.1.1.2. Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	1091
Nro. de Conexiones	1322
Densidad de la Red	0.00220990758415392
Modularidad	0.746
Núm. Comunidades	125

Tabla 1. Red 8M del 4-7 marzo - Resumen de pruebas estadísticas

4.1.1.3. Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	EsMomento_pe	91	-	-
2	duhnia	88	-	-
3	lamula	70	-	-
4	MarisaGlave	56	-	-
5	emmaylola	55	-	-
6	paromujeresperu	45	-	-
7	NUMSTLosOlivos	40	-	-
8	amnistiaperu	37	-	-
9	CMPFloraTristan	37	-	-
10	Actua_Pe	30	-	-
1	NUMSTLosOlivos	-	0.0012913331790173602	-
2	paromujeresperu	-	0.001029893972624026	-
3	Actua_Pe	-	0.0010020769806702527	-
4	CMPFloraTristan	-	0.0008714780621925468	-
5	paridadperu	-	0.00046225991597134187	-
6	DemusPeru	-	0.00038757312457618286	-
7	emmaylola	-	0.00022469622553025382	-
8	femimisha	-	0.00020509916673386232	-
9	NuevoPeruGP	-	0.00020372717203409444	-
10	promsex	-	0.00015414881970344857	-
1	amnistiaperu	-	-	0.3730404935191629
2	NUMSTLosOlivos	-	-	0.33545796110813436
3	paromujeresperu	-	-	0.3009461543087807

4	InfoLaPlaza	-	-	0.2780094797617801
5	Actua_Pe	-	-	0.2626336185890783
6	MadresAccionPe	-	-	0.25200572028229384
7	emmaylola	-	-	0.22541018092755843
8	NoticiasSER	-	-	0.19836199013289896
9	Graciela_Jimena	-	-	0.1964647080124585
10	CMPFloraTristan	-	-	0.18206943810530213

Tabla 2. Red 8M del 4-7 marzo - Nodos influyentes por tipo de Centralidad

4.1.1.4. Visualización y exploración de la Red

Luego que en Gephi se cargaran los datos desde los archivos de tipo CSV, uno para los nodos y el otro para los enlaces (generados durante la etapa de minería de datos), aplicamos dos de los diseños más populares: (1) Fruchterman-Reingold y (2) Force Atlas 2. En ambos diseños, se han coloreado los nodos en función de sus agrupaciones, según el resultado de la modularidad como parámetro de la partición, mientras que el tamaño de los nodos que refleja cuán importantes son estos para la red, se obtiene a partir de la centralidad Eigenvector.

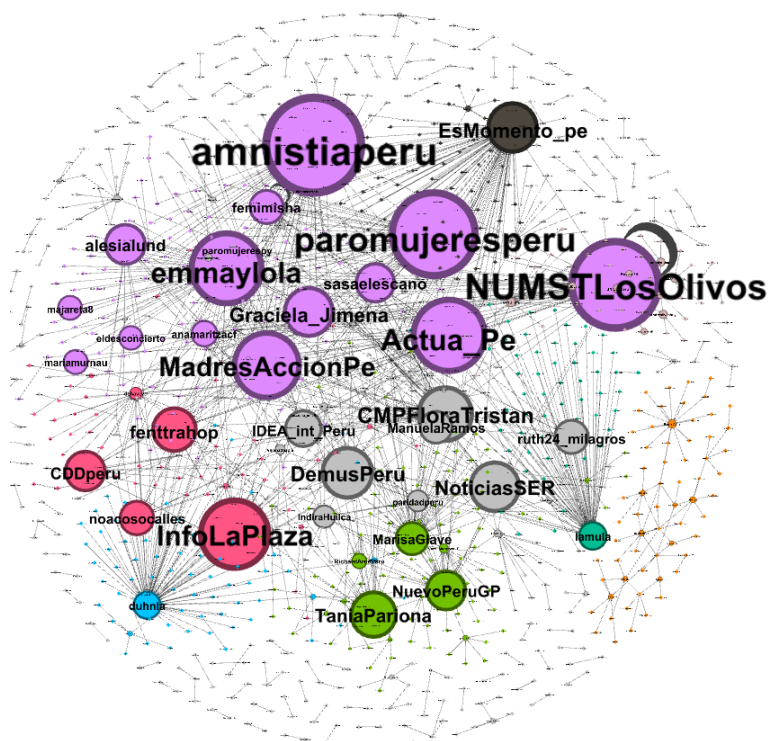


Figura 13. Diseño Fruchterman-Reingold para la Red 8M del 4-7 marzo

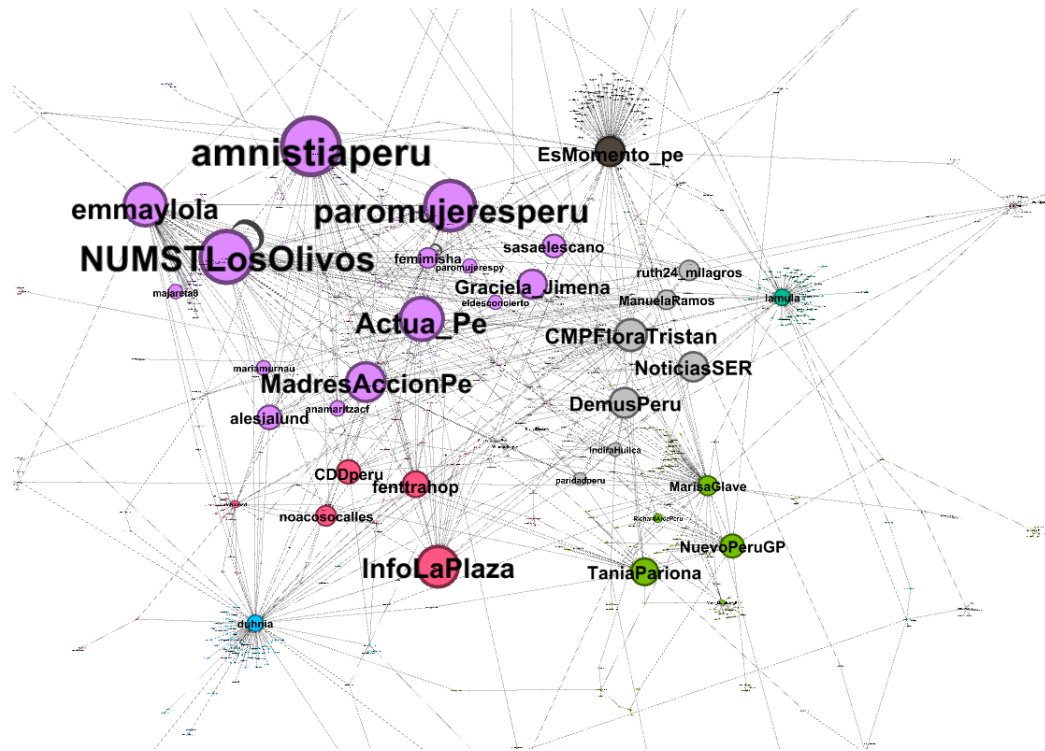


Figura 14. Diseño Force Atlas 2 para la Red 8M del 4-7 marzo

4.1.1.5. Principales Comunidades

Esta red de 1091 nodos y 1322 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 125 comunidades, de las que podemos destacar las siguientes:















#	Comunidad	Descripción	Subgrupos	% Participación
1	Política	Conformado por agrupaciones políticas, líderes de partidos políticos y parlamentarios.		10.17
2	Periodismo	Medio de comunicación virtuales, televisivos, escritos, etc.	 	5.41 5.77
3	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.	    	12.65 7.06 6.97 3.85 5.5
4	Organizaciones Deportivas / Culturales y Educativas	Actores, músicos, cantantes, clubes deportivos, etc.		4.22
5	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		38.40

Tabla 3. Red 8M del 4-7 marzo - Principales Comunidades

4.1.1.6. Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	amnistiaperu		Movimiento global independiente conformado por 7 millones de activistas en 150 países. Promueven y protegen los DD.HH en el mundo. https://twitter.com/amnstiaperu
2	NUMSTLosOlivos		Colectivo feminista ubicado en el distrito de Los Olivos, uno de los más poblados de Lima, Perú https://twitter.com/NUMSTLosOlivos
3	paromujeresperu		Plataforma Paro Internacional de Mujeres que visibiliza y lucha contra la desigualdad y la violencia hacia las mujeres. https://twitter.com/paromujeresperu
4	InfoLaPlaza		Periodistas independientes que no tienen reparo en enfrentarse al excluyente sistema económico ni al


			represor sistema político. https://twitter.com/InfoLaPlaza
5	Actua_Pe		Red de vigilancia, información y propuesta bajo un concepto lúdico y de interacción que conecta a personas y polític@s. https://twitter.com/Actua_Pe

Tabla 4. Red 8M del 4-7 marzo - Nodos influyentes

4.1.1.7. Observaciones

- En esta red existe predominancia de la Comunidad conformada por Organizaciones de Derechos Humanos y Feministas
- Los representantes de la Comunidad Política son parlamentarias (congresistas) pertenecientes a una de las agrupaciones políticas de la izquierda peruana.
- Es extraño que no exista participación del Gobierno, en especial del Ministerio de la Mujer y Poblaciones Vulnerables (que específicamente, el usuario en Twitter [@MimpPeru](https://twitter.com/MimpPeru) no figure en la red).

4.1.2 Grafo de la red 8M en el día central 8 marzo

4.1.2.1. Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	1521
Nro. de Conexiones	1487
Densidad de la Red	0.0012837814457247656
Modularidad	0.859
Núm. Comunidades	221

Tabla 5. Red 8M el 8 de marzo - Resumen de pruebas estadísticas

4.1.2.2. Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	pictoline	132	-	-
2	MaribelToledo	113	-	-
3	Maricarmen_Cix	57	-	-
4	discolocontilde	56	-	-
5	AlbertoBelaunde	46	-	-
6	amnistiaperu	26	-	-
7	EsMomento_pe	25	-	-
8	IndiraHuilca	25	-	-
9	Universitario	21	-	-
10	VeryNiceTrue	20	-	-
1	MidisPeru	-	3.768060704757285e-05	-
2	WaykaPeru	-	3.118395066006029e-05	-
3	NUMSTLosOlivos	-	2.9451508956723608e-05	-
4	AlbertoBelaunde	-	1.9273413949620597e-05	-
5	vesdape	-	1.6458196181698488e-05	-
6	EsMomento_pe	-	1.0178095007103011e-05	-
7	MIDIS_SISFOH	-	8.44565330376633e-06	-
8	MidisFoncodes	-	7.795987665015073e-06	-
9	rmapalacios	-	6.929766813346731e-06	-
10	NuevoPeruGP	-	5.630435535844219e-06	-
1	MidisPeru	-	-	0.5436886573651096
2	MidisQaliWarma	-	-	0.5436886570150531
3	MidisPAIS	-	-	0.4563106890770202
4	MIDIS_SISFOH	-	-	0.2955975493367796
5	MIDIS_CunaMas	-	-	0.2955975493367796
6	MidisJuntos	-	-	0.16071313974478685
7	Universitario	-	-	0.0008008604391637421
8	amnistiaperu	-	-	0.0002669620053462233
9	Graciela_Jimena	-	-	0.0002669571591057909
10	Actua_Pe	-	-	0.00026695552247675173

Tabla 6. Red 8M el 8 de marzo - Nodos influyentes por tipo de Centralidad

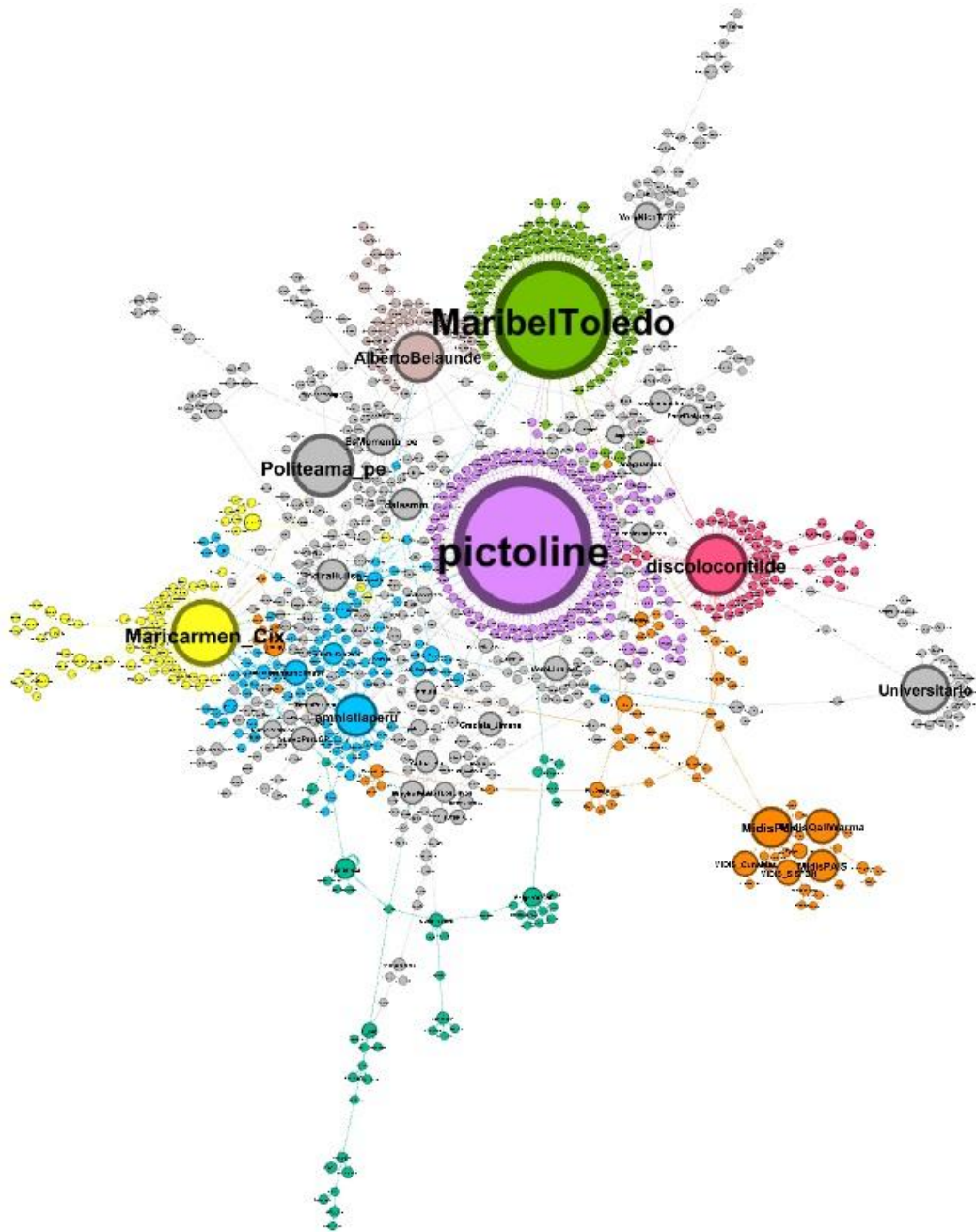


Figura 16. Diseño Force Atlas 2 para la Red 8M del 8 marzo





4.1.2.4. Principales Comunidades

Esta red de 1521 nodos y 1487 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 221 comunidades, de las que podemos destacar las siguientes:

#	Comunidad	Descripción	Subgrupos	% Participación
1	Política	Conformado por agrupaciones políticas, líderes de partidos políticos y parlamentarios.	 	2.96 3.09
2	Periodismo	Medio de comunicación virtuales, televisivos, escritos, etc.	   	8.68 4.80 6.64 3.16
3	Gobierno	Ministerios, Ministros de estado, Alcaldías, Poder Judicial, etc.		4.40
4	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.	 	5.98 4.01
5	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		56.28

Tabla 7. Red 8M el 8 de marzo - Principales Comunidades

4.1.2.5. Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	pictoline		Noticias e información en imágenes. https://twitter.com/pictoline
2	MaribelToledo		Maribel Toledo, periodista y consultora https://twitter.com/MaribelToledo
3	Maricarmen_Cix		Maricarmen Chinchay, periodista del diario La República https://twitter.com/Maricarmen_Cix
4	amnistiaperu		Movimiento global independiente conformado por 7 millones de activistas en 150 países. Promueven y protegen los DD.HH en el mundo. https://twitter.com/amnstiaperu


5	AlbertoBelaunde		Congresista de la Bancada Liberal, docente universitario. https://twitter.com/AlbertoBelaunde
6	MidisPeru		Ministerio de Desarrollo e Inclusión Social https://twitter.com/MidisPeru

Tabla 8. Red 8M el 8 de marzo - Nodos influyentes

4.1.2.6. Observaciones

- En esta red se observa que existen casi 50% más nodos que en la red 8M correspondiente al periodo del 4 al 7 de marzo. Así mismo, en esta fecha central de celebración 8M del día internacional de la mujer, participan entidades gubernamentales pertenecientes al Ministerio de Desarrollo e Inclusión Social (MIDIS) en color naranja.
- Durante el 8 de Marzo, Maribel Toledo (@MaribelToledo) y Maricarmen Chinchay (@Maricarmen_Cix), demuestran ser comunicadoras bastante influyentes (a pesar que no tienen participación en el 8M durante el periodo del 4 al 7 de marzo), y en menor medida las periodistas Verónica Linares (@VeroLinaresC) y Rosa María Palacios (@rmapalacios).
- Parlamentarios como Alberto de Belaunde (@AlbertoBelaunde) e Indira Huilca (@IndiraHuilca) destacan como políticos.
- El registro en fotografías las registra Pictoline (@pictoline), en la comunidad Cultural (color morado), dejando en evidencia, que muchas personas hacen seguimiento a las publicaciones gráficas de dicho usuario.
- Es extraño que el Ministerio de la Mujer y Poblaciones Vulnerables, con el usuario en Twitter [@MimpPeru](https://twitter.com/MimpPeru) no figure en la red. Se verá más adelante, en el apartado 4.2 Red del Día de la Mujer que hace una breve aparición.
- Sorprende que el diario El Comercio (@elcomercio_peru) no haya tenido relevante actividad en Twitter con el hashtag 8M.

4.1.3 Grafo de la red 8M del 9-12 marzo

4.1.3.1. Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	1368
Nro. de Conexiones	1188
Densidad de la Red	0.001270550186732376
Modularidad	0.918
Núm. Comunidades	254

Tabla 9. Red 8M del 9-12 marzo - Resumen de pruebas estadísticas

4.1.3.2. Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	Maricarmen_Cix	88	-	-
2	pictoline	43	-	-
3	MaribelToledo	42	-	-
4	discolocontilde	38	-	-
5	1alos22	36	-	-
6	NuevoPeruMov	25	-	-
7	MarisaGlave	22	-	-
8	QuintoPoderPe	21	-	-
9	EsMomento_pe	20	-	-
10	arlettecontrers	20	-	-
1	NUMSTLosOlivos	-	4.444868105233056e-05	-
2	paromujeresperu	-	2.570526133146827e-05	-
3	femimisha	-	8.568420443822758e-06	-
4	emmaylola	-	5.890789055128146e-06	-
5	mmolerosuarez	-	5.355262777389224e-06	-
6	CONFIEP	-	3.7486839441724563e-06	-
7	claudiacisneros	-	2.677631388694612e-06	-
8	ray_asm	-	1.0710525554778447e-06	-
9	bobsparz	-	1.0710525554778447e-06	-
10	MIDIS_SISFO	-	5.355262777389223e-07	-
1	NUMSTLosOlivos	-	-	0.42810892094881575
2	InfoLaPlaza	-	-	0.4280416424613831
3	celestemacdou	-	-	0.4280409323981912
4	_lamoreau	-	-	0.4280402223349992
5	8Mdublin	-	-	0.4280402223349992
6	UNICEFperu	-	-	0.20475250839945605
7	Universitario	-	-	0.20475250839945575
8	Gender_Lab	-	-	0.0001090124515364885
9	sedoac	-	-	4.421918527558914e-05
10	RebecaLane6	-	-	4.421918527558914e-05

Tabla 10. Red 8M del 9-12 marzo - Nodos influyentes por tipo de Centralidad

4.1.3.3. Visualización y exploración de la Red

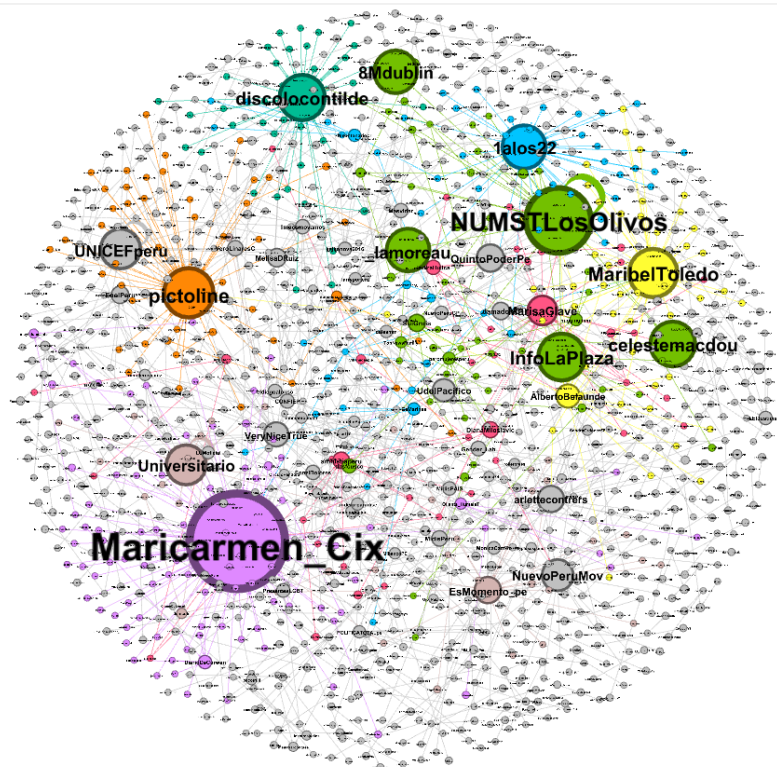


Figura 17. Diseño Fruchterman-Reingold para la Red 8M del 9-12 marzo

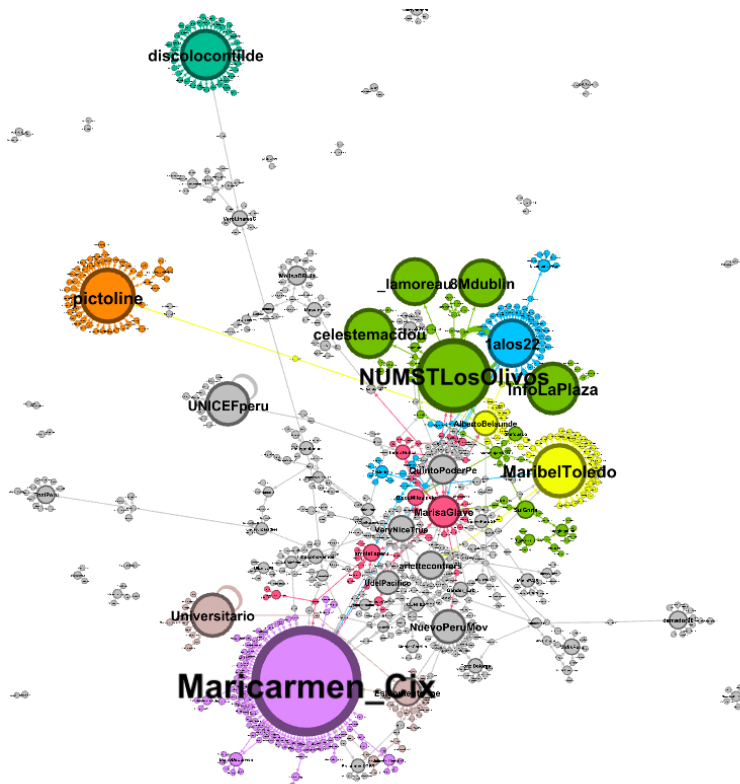


Figura 18. Diseño Force Atlas 2 para la Red 8M del 9-12 marzo

4.1.3.4. Principales Comunidades

Esta red de 1368 nodos y 1188 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 254 comunidades, de las que podemos destacar las siguientes:





#	Comunidad	Descripción	Subgrupos	% Participación
1	Política	Conformado por agrupaciones políticas, líderes de partidos políticos y parlamentarios.		2.78
				3.00
2	Periodismo	Medio de comunicación virtuales, televisivos, escritos, etc.		7.09
				3.80
				2.85
				3.73
3	Gobierno	Ministerios, Ministros de estado, Alcaldías, Poder Judicial, etc.		2.70
4	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.		4.09
5	Organizaciones Deportivas / Culturales	Actores, músicos, cantantes, clubes deportivos, etc.		3.95
6	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		66.01

Tabla 11. Red 8M del 9-12 marzo - Principales Comunidades

4.1.3.5. Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	NUMSTLosOlivos		Colectivo feminista ubicado en el distrito de Los Olivos, uno de los más poblados de Lima, Perú https://twitter.com/NUMSTLosOlivos
2	MaribelToledo		Maribel Toledo, periodista y consultora https://twitter.com/MaribelToledo
3	Maricarmen_Cix		Maricarmen Chinchay, periodista del diario La República https://twitter.com/Maricarmen_Cix




4	InfoLaPlaza		Periodistas independientes que no tienen reparo en enfrentarse al excluyente sistema económico ni al represor sistema político. https://twitter.com/InfoLaPlaza
5	pictoline		Noticias e información en imágenes. https://twitter.com/pictoline
6	UNICEFperu		Promoción, protección y cumplimiento de los derechos de las niñas, niños y adolescentes. https://twitter.com/UNICEFperu

Tabla 12. Red 8M del 9-12 marzo - Nodos influyentes

4.1.3.6. Observaciones

- Luego de la fecha central del 8M la red se dispersa, pero las comunicadoras Maribel Toledo y Maricarmen Chinchay aun predominan como personas influyentes en la red.
- Se advierte también la interacción entre políticos como Marisa Glave, Indira Huilca, Yeni Vilcatoma y Alberto de Belaunde, así como la escasa participación del gobierno a través del MIDIS. Se visualiza además la presencia de Arlette Contreras (@arlettecontrers), abogada y activista peruana que fue víctima hace unos años de violencia contra la mujer y que a la fecha continua buscando justicia.

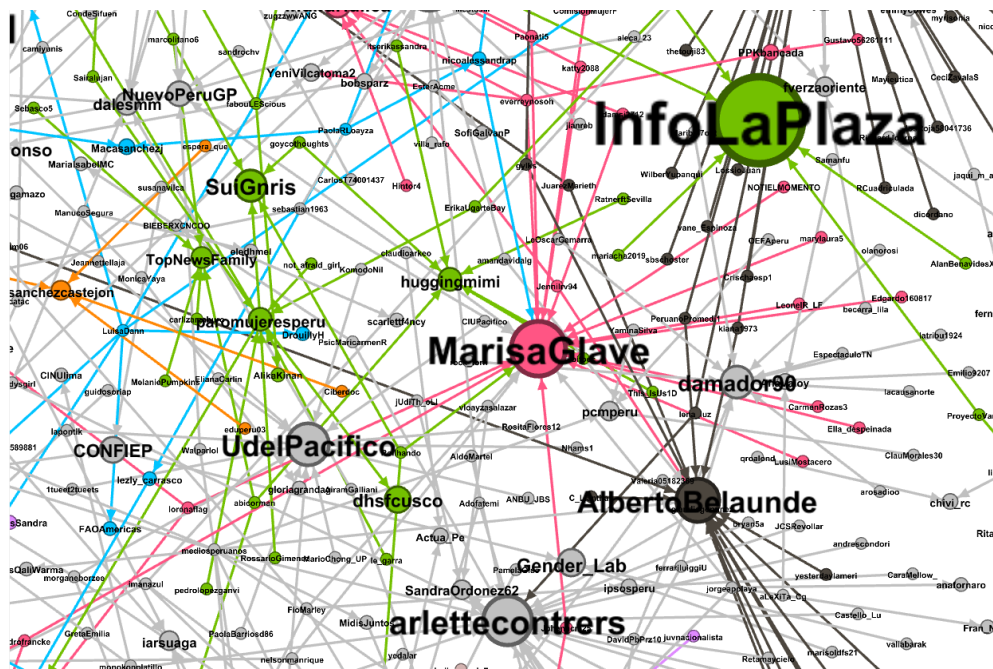


Figura 19. Participación de políticos en la red 8M

4.2 Red Día de la Mujer

Esta red complementa a la Red 8M y considera los hashtags #Diadelamujer y #DiaInternacionaldeLaMujer.

4.2.1 Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	5807
Nro. de Conexiones	7531
Densidad de la Red	0.0004443067956813379
Modularidad	0.780
Núm. Comunidades	475

Tabla 13. Red Día de la Mujer - Resumen de pruebas estadísticas

4.2.2 Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	Pictoline	278	-	-
2	saldelsol	183	-	-
3	KeikoFujimori	137	-	-
4	HistoriaEnFotos	130	-	-
5	ONU_es	119	-	-
6	rosamariabartra	114	-	-
7	ClubALoficial	96	-	-
8	Universitario	89	-	-
9	BoseOficial	87	-	-
10	minieides	85	-	-
1	UNICEFperu	-	0.0002863828828949114	-
2	CINULima	-	0.00019155686460559522	-
3	valiabarak	-	0.00018961247628731409	-
4	CAREenPeru	-	0.00015021537117097176	-
5	MidisPeru	-	0.0001452505546105591	-
6	MidisFoncodes	-	0.00010241170414954818	-
7	monich16	-	9.346118823884407e-05	-
8	noticias_tvperu	-	9.016779398661814e-05	-
9	tvperupe	-	8.555407501165298e-05	-
10	MovistarPeru	-	8.22755158686713e-05	-
1	Pictoline	-	-	1.0
2	KeikoFujimori	-	-	0.6600028301308287
3	saldelsol	-	-	0.6512560614956374
4	ONU_es	-	-	0.63028945860446
5	LUMoficial	-	-	0.5294689212789754
6	ONUMujeres	-	-	0.5168365767938048
7	HistoriaEnFotos	-	-	0.4642619263870845
8	Universitario	-	-	0.4450007904528109
9	rosamariabartra	-	-	0.4328182804444491
10	CINULima	-	-	0.4074195488431143

Tabla 14. Red Día de la Mujer - Nodos influyentes por tipo de Centralidad

4.2.4 Principales Comunidades

Esta red de 5807 nodos y 7531 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 476 comunidades, de las que podemos destacar las siguientes:








#	Comunidad	Descripción	Subgrupos	% Participación
1	Política	Conformado por agrupaciones políticas, líderes de partidos políticos, parlamentarios y analistas políticos	  	7.90 3.70 3.79
2	Periodismo	Medios de comunicación virtuales, televisivos, escritos, etc.	 	4.86 3.75
3	Gobierno	Ministerios, Ministros de Estado, Alcaldías, Poder Judicial, etc.	 	3.38 4.51
4	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.		5.67
5	Organizaciones Deportivas / Culturales	Actores, músicos, cantantes, clubes deportivos, etc.		3.91
6	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		58.53

Tabla 15. Red Día de la Mujer - Principales Comunidades

4.2.5 Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	pictoline		Noticias e información en imágenes. https://twitter.com/pictoline
2	KeikoFujimori		Fundadora y Presidenta de @PFuerzaPopular. https://twitter.com/KeikoFujimori
3	saldelsol		Presidente del Consejo de Ministros del Perú. https://twitter.com/saldelsol
4	ONU_es		Cuenta oficial de #NacionesUnidas en español. https://twitter.com/Maricarmen_Cix
5	LUMoficial		El Lugar de la Memoria (LUM) es un museo sobre el período de violencia 1980-2000 en el Perú, desde el enfoque de los derechos humanos.


			https://twitter.com/LUMoficial
6	ONUMujeres		Agencia de las Naciones Unidas para la igualdad de género y el empoderamiento de las mujeres. https://twitter.com/ONUMujeres

Tabla 16. Red Día de la Mujer - Nodos influyentes

4.2.6 Observaciones

- Esta red es mucho más grande que la red del 8M (casi la cuadruplica). La posible causa, podría deberse a la mayor participación de la Comunidad Periodismo dentro de esta red, y su menor figuración dentro de la red del 8M. Se demostraría que la Comunidad Periodismo es importante en cuanto al tamaño de las redes, y que la cobertura periodística puede lograr de que las redes sean más grandes.
- Se incorporan nuevos actores o nodos de relevancia como Keiko Fujimori y Rosa María Bartra a nivel político en comunidad de Política de color morado.
- La red de Keiko Fujimori está, integrada por nodos que comparten su nombre como @RedesKeiko, @JovenesKeiko y @KeikoTrabaja las cuales guardan relaciones con otros nodos de mayor influencia.
- Las comunicadoras Maribel Toledo y Maricarmen Chinchay, que fueran ambas de gran influencia en la red 8M ya no participan en esta red del DiadelaMujer.
- Esta red es más política, de entidades gubernamentales, de medios de comunicación, de organizaciones de defensa de los derechos, actores y cantantes en vez que de colectivos feministas y/o de defensa de los derechos de las mujeres. Sus principales participantes pertenecen a comunidades de orden político, gubernamental, cultural y medios de comunicación.

- El Ministerio de la Mujer y Poblaciones Vulnerables ([@MimpPeru](#)) hace su aparición en esta red solo con tres relaciones como se muestra en la figura 22.

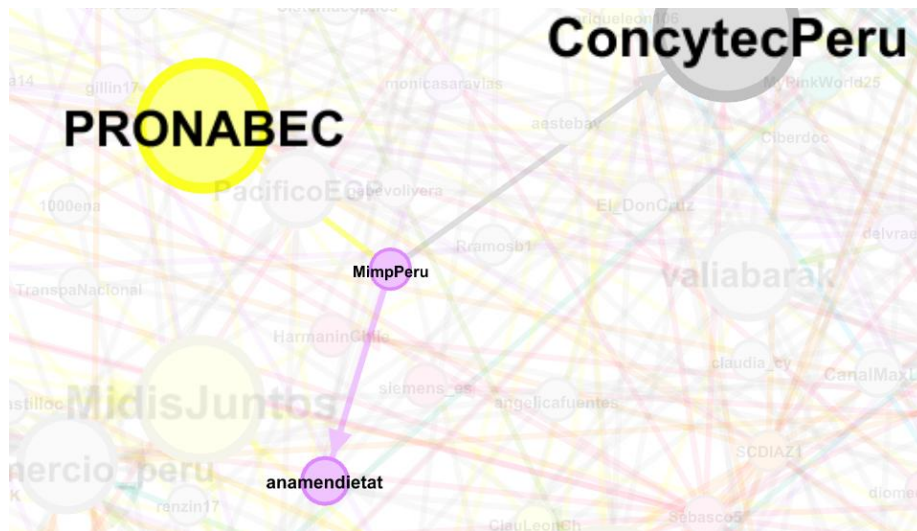


Figura 22. Interacciones del MIMP en el Día Internacional de la Mujer

4.3 Red Marcha 8M

Esta red surge como una actividad durante el evento de conmemoración del Día internacional de la Mujer. Hashtag #Marcha8M

4.3.1 Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	136
Nro. de Conexiones	143
Densidad de la Red	0.015468409586056645
Modularidad	0.744
Núm. Comunidades	21

Tabla 17. Red Marcha 8M - Resumen de pruebas estadísticas

4.3.2 Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	ManuelaRamos	21	-	-
2	amniaperu	19	-	-
3	CMPFloraTristan	12	-	-
4	ruth24_milagros	10	-	-
5	lorenaflag	8	-	-
6	elcomercio_peru	7	-	-
7	Graciela_Jimena	7	-	-
8	JaquiNavidad	7	-	-
9	paridadperu	7	-	-
10	Vaniadiana	6	-	-
1	ruth24_milagros	-	0.00845771144278607	-
2	ManuelaRamos	-	0.007435046987285794	-
3	paridadperu	-	0.005362078496406855	-
4	CMPFloraTristan	-	0.0029021558872305143	-
5	lorenaflag	-	0.002211166390270868	-
6	elcomercio_peru	-	0.0003316749585406302	-
7	CalandriaPeru	-	0.0	-
8	fatimavaldivia	-	0.0	-
9	_believe1108	-	0.0	-
10	sebacuchufli	-	0.0	-
1	ManuelaRamos	-	-	0.5951801751507403
2	ruth24_milagros	-	-	0.4815108764021848
3	Villanadeapie	-	-	0.3678415780304207
4	amniaperu	-	-	0.2975900876238146
5	paridadperu	-	-	0.297590087566399
6	CMPFloraTristan	-	-	0.18392078881797166
7	Vaniadiana	-	-	0.18392078880361765
8	CarolRuizMoran	-	-	0.18392078878567522
9	Lima_ECpe	-	-	0.00034402788005233266
10	larepublica_pe	-	-	2.1184910953931454e-10

Tabla 18. Red Marcha 8M - Nodos influyentes por tipo de Centralidad

4.3.3 Visualización y exploración de la Red

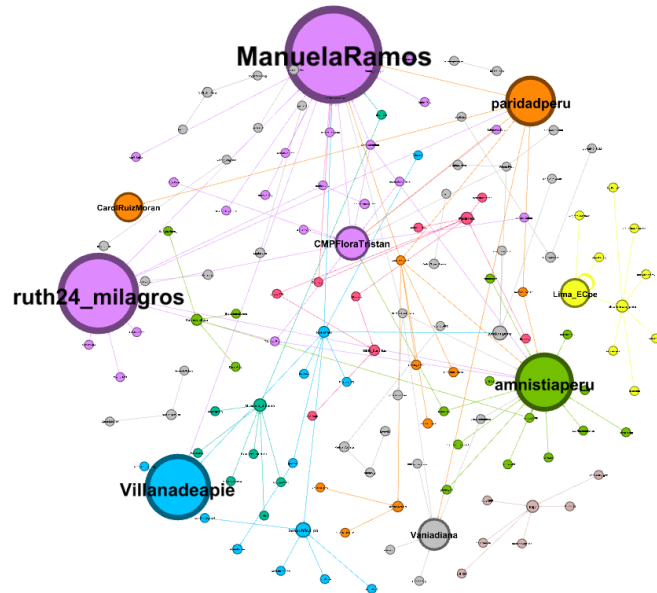


Figura 23. Diseño Fruchterman-Reingold para la Red Marcha 8M

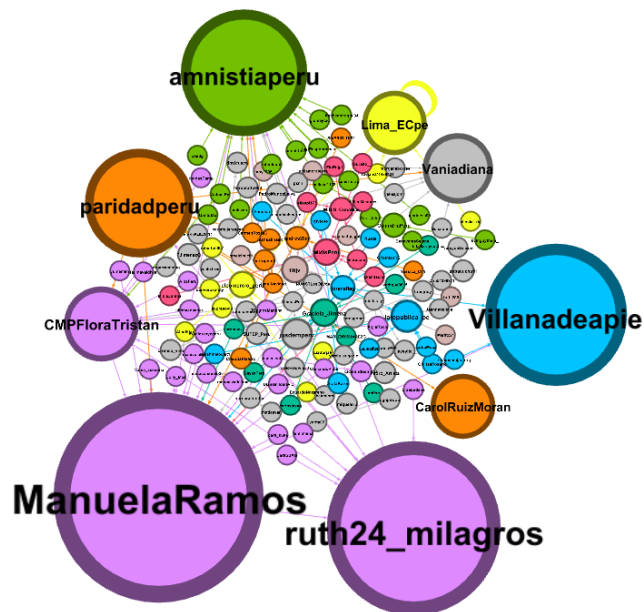


Figura 24. Diseño Force Atlas 2 para la Red Marcha 8M

4.3.4 Principales Comunidades

Esta red de 136 nodos y 143 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 21 comunidades, de las que podemos destacar las siguientes:














#	Comunidad	Descripción	Subgrupos	% Participación
1	Periodismo	Medios de comunicación virtuales, televisivos, escritos, etc.	  	7.35 9.56 5.88
2	Gobierno	Ministerios, Ministros de Estado, Alcaldías, Poder Judicial, etc.		6.62
3	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.	   	17.65 7.35 12.50 5.15
4	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		27.94

Tabla 19. Red Marcha 8M - Principales Comunidades

4.3.5 Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	ManuelaRamos		Organización feminista peruana comprometida con la igualdad, el respeto a la diversidad y el ejercicio autónomo de los derechos de las mujeres. Fundada en 1978. https://twitter.com/ManuelaRamos
2	Villanadeapie		Periodista. Feminista. Co-fundador de @WaykaPeru. Comunicadora en Flora Tristán / Manuela Ramos. https://twitter.com/Villanadeapie
3	amnistiaperu		Movimiento global independiente conformado por 7 millones de activistas en 150 países. Promueven y protegen los DD.HH en el mundo. https://twitter.com/amnstiaperu
4	paridadperu		Movimiento feminista que apuesta por una democracia con paridad para garantizar el derecho de la mujer a la participación política. https://twitter.com/paridadperu
5	CMPFloraTristan		Institución feminista creada en 1979. Su misión es combatir las causas estructurales que restringen y afectan la ciudadanía de las mujeres

			peruanas. https://twitter.com/CMPFloraTristan
--	--	--	--

Tabla 20. Red Marcha 8M - Nodos influyentes

4.3.6 Observaciones

- La red conformada para la Marcha del 8M fue pequeña y no generó gran cantidad de relaciones. Básicamente, dicha marcha concentró a personas que participan en movimientos feministas.
- Se distingue al diario La Republica como un medio de comunicación escrita que interactuó más que otros medios (diario El Comercio y WaykaPeru).
- La participación de entidades gubernamentales es casi nula o imperceptible, como es el caso del MIDIS.

4.4 Red de Acoso

Esta red representa a las personas e instituciones/organizaciones que se relacionan a través del tema de acoso hacia la mujer, tan importante hoy día en el Perú.

4.4.1 Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	4805
Nro. de Conexiones	10920
Densidad de la Red	0.000944408968939342
Modularidad	0.555
Núm. Comunidades	157

Tabla 21. Red de Acoso - Resumen de pruebas estadísticas

4.4.2 Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	rmapalacios	222	-	-
2	piacostar	186	-	-
3	yonhy_lescano	186	-	-
4	IndiraHuilca	184	-	-
5	saldelsol	184	-	-
6	noakeikope	171	-	-
7	SolCn	169	-	-
8	RPPNoticias	140	-	-
9	Maricarmen_Cix	139	-	-
10	RecuerdosPeru	136	-	-
1	promsex	-	0.004063067459178799	-
2	_zavandija	-	0.0029127799918107836	-
3	rmapalacios	-	0.002212176764233241	-
4	PacitaGuerra	-	0.0018289665907613639	-
5	paridadperu	-	0.0017455592623667746	-
6	Josefina_28	-	0.0016045632930876126	-
7	OZapat1	-	0.0014955293631850623	-
8	NuevoPeruGP	-	0.0014417215566216411	-
9	IndiraHuilca	-	0.0014206513112406123	-
10	Heli21270798	-	0.0014107780454180843	-
1	IndiraHuilca	-	-	0.30962810103900634
2	Jeannettellaja	-	-	0.26549071765092275
3	dchavezd	-	-	0.22875640371440914
4	TaniaPariona	-	-	0.21340185188860422
5	Vero_Mendoza_F	-	-	0.18670793441487668
6	luis_francia	-	-	0.18637643655620237
7	mmolerosuarez	-	-	0.18413622934770685
8	NuevoPeruGP	-	-	0.18345333297018915
9	promsex	-	-	0.16990522822766124
10	radio_santarosa	-	-	0.1675268562238596

Tabla 22. Red de Acoso - Nodos influyentes por tipo de Centralidad

4.4.3 Visualización y exploración de la Red

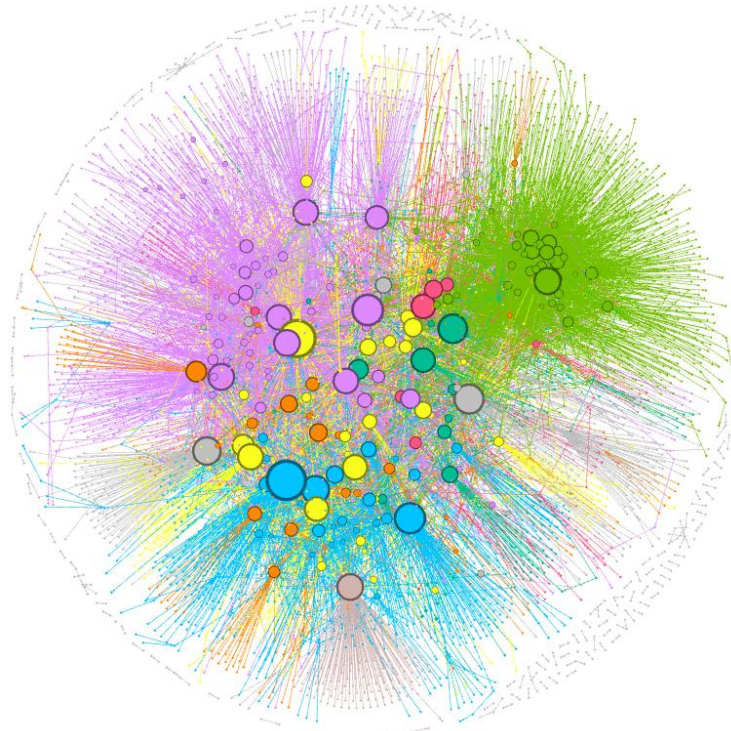


Figura 25. Diseño Fruchterman-Reingold para la Red Acoso en Perú

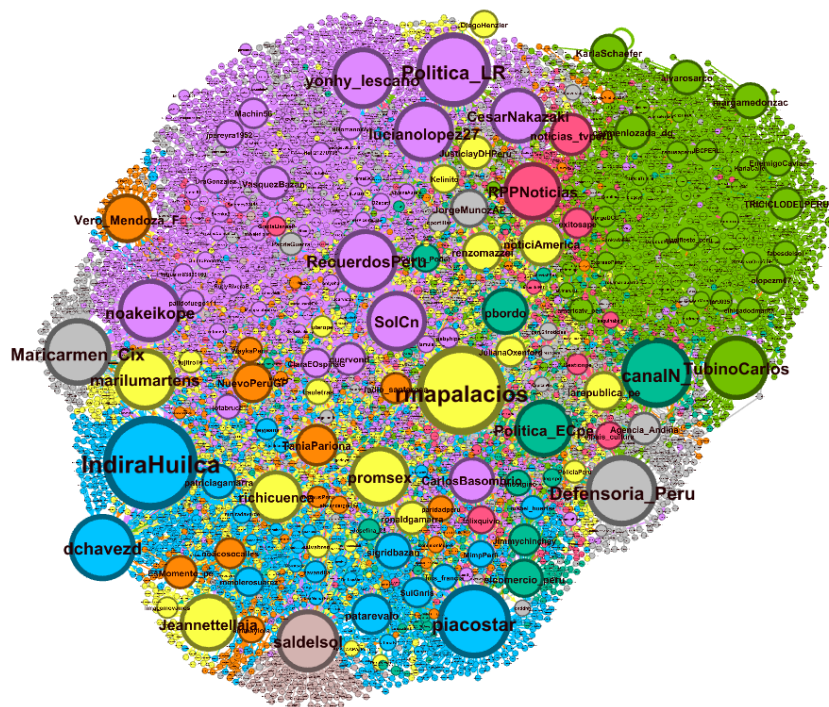


Figura 26. Diseño Force Atlas 2 para la Red Acoso en Perú

4.4.4 Principales Comunidades

Esta red de 4805 nodos y 10920 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 157 comunidades, de las que podemos destacar las siguientes:








#	Comunidad	Descripción	Subgrupos	% Participación
1	Política	Conformado por agrupaciones políticas, líderes de partidos políticos, parlamentarios y analistas políticos		19.54
2	Periodismo	Medios de comunicación virtuales, televisivos, escritos, etc.	   	8.78 5.04 4.64 4.39
3	Gobierno	Ministerios, Ministros de estado, Alcaldías, Poder Judicial, etc.	 	2.41 6.04
4	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.	 	23.91 16.13
6	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		9.12

Tabla 23. Red de Acoso - Principales Comunidades

4.4.5 Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	IndiraHuilca		Feminista y socialista. Congresista de la República por Lima Metropolitana. Socióloga por la UNMSM. Militante del Movimiento Nuevo Perú. https://twitter.com/IndiraHuilca
2	rmapalacios		Abogada. Periodista expulsada por ser excesivamente independiente (Mario Vargas Llosa). https://twitter.com/rmapalacios
3	yonhy_lescano		Congresista de la República por Lima Metropolitana. Militante del Partido Político Acción Popular. https://twitter.com/yonhy_lescano
4	promsex		Centro de Promoción y Defensa de los Derechos Sexuales y Reproductivos, Perú. https://twitter.com/promsex
5	Vero_Mendoza_F		Cusqueña, mamá, antropóloga.

			Presidenta del @Mov_NuevoPeru (2018-2019). https://twitter.com/Vero_Mendoza_F
--	--	--	---

Tabla 24. Red de Acoso - Nodos influyentes

4.4.6 Observaciones

- La fortaleza de esta red radica en los nodos que pertenecen a la Comunidad compuesta por Organizaciones de Derechos Humanos y Feministas.
- Si bien, las comunidades de Política y Gobierno tienen participación activa, algunos políticos y/o entidades de gobierno figurarían como pertenecientes a la comunidad compuesta por Organizaciones de Derechos Humanos y Feministas. Caso de la señora Indira Huilca, Congresista de la Republica, en esta red su papel pareciera ser más activo como feminista que como política, al igual que Verónica Mendoza. Del mismo modo, el Ministerio de la Mujer y Poblaciones Vulnerables, antes de figurar como entidad de la Comunidad de Gobierno, se confunde también como una entidad dentro de la lucha de Derechos Humanos y Feministas.
- El congresista Yonhy Lescano ([@yonhy_lescano](#)) que aparece en el tercer lugar dentro de los actores influyentes en la red de Acoso, no necesariamente lo hace como defensor de los derechos de las mujeres, más bien, su aparición es de relevancia debido a que previa a la fecha del 8 de marzo, específicamente el 4 de marzo último, fue denunciado por el delito de acoso sexual por una periodista radial. Debido a esta denuncia, fue separado de su cargo de Congresista de la Republica por un período de 120 días.

4.5 Red de Femicidio

Esta red representa a las personas e instituciones/organizaciones relacionadas con el tema de Femicidio.

4.5.1 Resumen de pruebas estadísticas

Tipo de red	Dirigida
Nro. de Nodos	625
Nro. de Conexiones	594
Densidad de la Red	0.003041025641025641
Modularidad	0.833
Núm. Comunidades	69

Tabla 25. Red de Femicidio - Resumen de pruebas estadísticas

4.5.2 Nodos influyentes por tipo de Centralidad

#	Nodo	Tipo de Centralidad		
		Degree	Betweenness	Eigenvector
1	saldelsol	184	-	-
2	mabel_huertas	40	-	-
3	elcomercio_peru	26	-	-
4	FiscaliaPeru	25	-	-
5	CNNEE	25	-	-
6	larepublica_pe	17	-	-
7	pucp	14	-	-
8	CINULima	13	-	-
9	Caretas	13	-	-
10	emmaylola	12	-	-
1	elcomercio_peru	-	0.00018006338231057334	-
2	Ana_Monzon_	-	6.688068485821296e-05	-
3	emmaylola	-	2.8295674363090095e-05	-
4	PanamericanaTV	-	1.2861670165040952e-05	-
5	OswaldoMolinaC	-	2.5723340330081904e-06	-
6	MAGDALENADELAFU	-	0.0	-
7	Lima_Cero	-	0.0	-
8	Tifannytaipe27	-	0.0	-
9	NoEstasSolaPeru	-	0.0	-
10	CoachingAhoraAL	-	0.0	-
1	PanamericanaTV	-	-	0.5547001419496722
2	BDPTV	-	-	0.5547001254183197
3	24HorasTV	-	-	0.5547001254183197
4	HRomanoWy	-	-	0.27735007097483605
5	mabel_huertas	-	-	0.00047197837837738207
6	Ana_Monzon_	-	-	6.014106024542193e-05
7	PeruECpe	-	-	5.993441833968587e-05
8	saldelsol	-	-	3.8030376331661384e-05
9	ManuelaRamos	-	-	2.7698281044860952e-05
10	UdelPacífico	-	-	5.794239036839012e-06

Tabla 26. Red de Femicidio - Nodos influyentes por tipo de Centralidad

4.5.3 Visualización y exploración de la Red

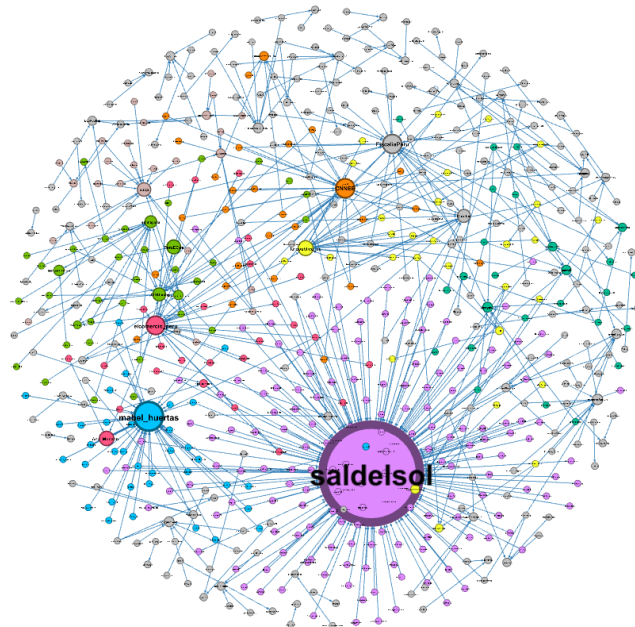


Figura 27. Diseño Fruchterman-Reingold para la Red Femicidio en Perú

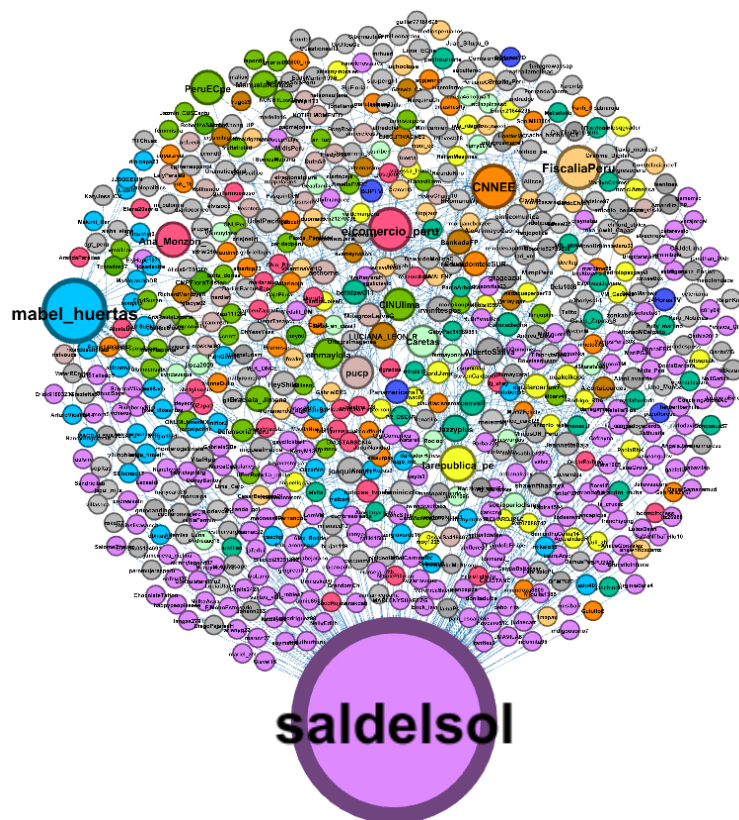


Figura 28. Diseño Force Atlas 2 para la Red Femicidio en Perú

4.5.4 Principales Comunidades

Esta red de 625 nodos y 594 conexiones representa a personas/organizaciones peruanas que se relacionan conformando 69 comunidades, de las que podemos destacar las siguientes:






#	Comunidad	Descripción	Subgrupos	% Participación
1	Política	Conformado por agrupaciones políticas, líderes de partidos políticos, parlamentarios y analistas políticos	 	4.16 2.08
2	Periodismo	Medios de comunicación virtuales, televisivos, escritos, etc.	     	4.32 5.92 4.80 4.64 3.04 0.80
3	Gobierno	Ministerios, Ministros de Estado, Alcaldías, Poder Judicial, Policía, etc.	 	27.20 4.00
4	Organizaciones de Derechos Humanos y Feministas	Organizaciones y personas del sector feminista identificado en la lucha por los derechos de las mujeres y niños.		6.08
5	Organizaciones Deportivas / Culturales y Educativas	Actores, músicos, cantantes, clubes deportivos, etc.		4.16
6	Otros	Otras personas o grupos que no alcanzan el 3% de representatividad.		28.8

Tabla 27. Red de Femicidio - Principales Comunidades

4.5.5 Nodos influyentes

Ítem	Nombre Nodo	Comunidad	Descripción
1	saldelsol		Presidente del Consejo de Ministros del Perú. https://twitter.com/saldelsol
2	mabel_huertas		Periodista. Todas las mañanas en Buenos Días Perú @BDPTV por @PanamericanaTV https://twitter.com/mabel_huertas
3	PeruECpe		Cuenta oficial de la sección Perú del diario El Comercio. https://twitter.com/PeruECpe




4	PanamericanaTV		Pionero de la televisión Peruana. Primeros en la información con programas @BDPTV , @24HorasTV y @PanoramaPTV . https://twitter.com/PanamericanaTV
5	Ana_Monzon_		Periodista de el diario El Comercio @elcomercio_peru https://twitter.com/Ana_Monzon_
6	ManuelaRamos		Organización feminista peruana comprometida con la igualdad, el respeto a la diversidad y el ejercicio autónomo de los derechos de las mujeres. Fundada en 1978. https://twitter.com/ManuelaRamos

Tabla 28. Red de Femicidio - Nodos influyentes

4.5.6 Observaciones

- Se observa que las principales comunidades que interactúan sobre el tema del Femicidio en Perú son aquellas relacionadas al Gobierno, al Periodismo y a la Política. En cambio, la comunidad que agrupa a las Organizaciones de Derechos Humanos y Feministas no es muy representativa, así como tampoco lo son aquellos miembros que no han sido clasificados dentro de alguna de estas comunidades.
- En la visualización de la red se distingue un gran nodo de color morado correspondiente al Primer Ministro de Perú Salvador del Solar ([@saldelsol](#)). En primera instancia, se podría pensar que el Primer Ministro lleva el problema del Femicidio sobre sus hombros, sin embargo, no se puede dar por cierta dicha afirmación, dado que la fecha de su nombramiento en el cargo fue el 11 de Marzo del 2019, por lo que se requiere analizar su participación durante otros periodos de tiempo.
- Es clara la participación de los principales medios de comunicación, las noticias sobre los femicidios acaparan las portadas de los diarios, los programas de noticias televisivos, radiales, agencias internacionales de noticias y son los principales periodistas los encargados de su difusión.
- En esta red, NO resalta a simple vista la participación de la Defensoría del Pueblo ([@Defensoria_Peru](#)), organismo constitucionalmente autónomo del Estado Peruano encargado de “defender y promover los derechos de las personas, supervisar la eficacia de la actuación de la administración estatal y la adecuada prestación de los servicios públicos para que todas las peruanas y peruanos, especialmente los que se encuentran en situación de vulnerabilidad, tengan una vida próspera y plena” [22]. El MIMP encargado de velar por las mujeres y las poblaciones vulnerables, es otro organismo del Estado que no destaca en esta red. Quien sí lo hace es la Fiscalía de la Nación ([@Fiscalia_Peru](#)), organismo que interactúa con la Defensoría del Pueblo y el MIMP. En la figura 29 se observa que la Fiscalía interactúa con la Defensoría, mas “no” con el MIMP.

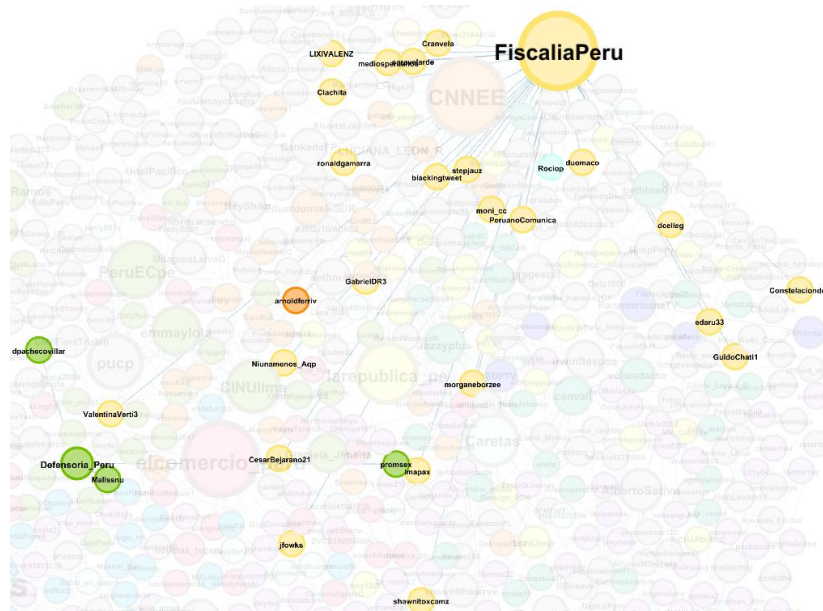


Figura 29. Enlaces de la Fiscalía Perú dentro de la Red de Femicidio

- En la figura 30, se visualiza que el MIMP ([@MimpPeru](#)), en color plomo, interactúa con la Fiscalía (@FiscaliaPeru) a través de la ComisionMujerF así como con la Defensoría del Pueblo (Defensoria_Peru).

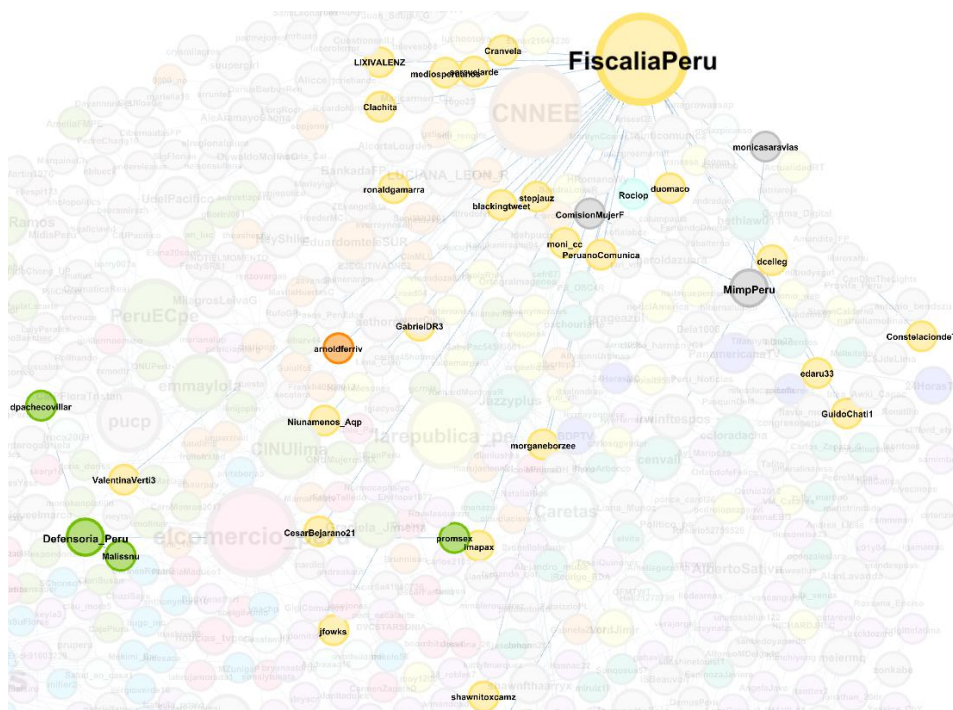


Figura 30. Enlaces del MIMP dentro de la Red de Femicidio

- No se ha encontrado ninguna relación ni con el Ministerio del Interior (MININTER) ni con la Policía Nacional del Perú (PNP).

5. Análisis y Observaciones

En el apartado cuatro se desarrollaron los modelos de cada una de las redes propuestas, calculando las principales estadísticas y señalando las principales observaciones. Así mismo, en base a los resultados gráficos obtenidos, se realizaron tres tipos de análisis de forma individual por cada red:

- **Análisis Estructural.** A efectos de conocer la estructura general de cada red, se evaluó el indicador de densidad. Esta medida indica la proporción de vínculos presentes entre los actores, frente al número total de conexiones posibles; tomando valores entre cero y uno.
- **Análisis de Centralidad del actor.** Para determinar el actor más importante de la red, se obtuvieron las mediciones de los siguientes indicadores: 1) centralidad de grado; 2) centralidad de intermediación y 3) centralidad de vector propio.
- **Análisis por Comunidades.** Permite reconocer patrones en los enlaces entre actores.

Las redes sociales modeladas fueron:

Red 1: 8 de Marzo – #8M

Red 2: Día de la Mujer - #DiaMujer

Red 3: Marcha 8M – #Marcha8M

Red 4: Acoso

Red 5: Femicidio

A continuación, se realiza un análisis estructural y de comunidades de forma comparativa para estas cinco redes. Así mismo, se realiza el análisis textual de las redes de Acoso y Femicidio a través de la visualización de nubes de palabras (Word Clouds).

5.1 Análisis Estructural de las Redes

Para poder calificar si la densidad es alta, media o baja, se ha adoptado la escala de medición expuesta en la tabla 29, la cual se basa en la proporción [23]

<i>Intervalos de valor</i>	<i>Categorías de la Densidad</i>
0.80 - 1	Muy Alta
0.60 - 0.79	Alta
0.40 – 0.59	Media
0.20 - 0.39	Baja
0 – 0.19	Muy Baja

Tabla 29. Escala para la medición de la Densidad de Red

En la tabla 30 se visualiza las medidas de densidad obtenidas para cada red así como su categorización:

Redes	Densidad	#Nodos	#Enlaces	Escala de densidad
Red 1: 8M (día central)	0.0012837814457247656	1521	1487	Muy baja
Red 2: DiaDeLaMujer	0.0004443067956813379	5807	7531	Muy baja
Red 3: Marcha 8M	0.015468409586056645	136	143	Muy baja
Red 4: Acoso	0.000944408968939342	4805	10920	Muy baja
Red 5: Femicidio	0.003041025641025641	625	594	Muy baja

Tabla 30. Categorización de la densidad por Red

Se puede observar que todas las redes presentan una densidad muy baja, esto significa que:

Red 1: 8M (día central): entre los 1521 actores que componen la misma, se dan el 0.1% de los vínculos posibles.

Red 2: DiaDeLaMujer: entre los 5807 actores que componen la misma, se dan el 0.04% de los vínculos posibles.

Red 3: Marcha 8M: entre los 136 actores que componen la misma, se dan el 1.5% de los vínculos posibles.

Red 4: Acoso: entre los 4805 actores que componen la misma, se dan el 0.09% de los vínculos posibles.

Red 5: Femicidio: entre los 625 actores que componen la misma, se dan el 0.3% de los vínculos posibles.

Al constatar que todas las redes poseen una densidad baja, esta situación indica que se trata de redes poco cohesionadas.

5.2 Análisis por Comunidades

A continuación, se describe y grafica cómo la participación de las comunidades incide de forma general en cada red. Con este análisis, se pretende determinar por Comunidad, que tan importante se considera la defensa de los derechos de la mujer peruana. En la tabla 31 se resume la medición de modularidad por cada red.

Redes	Modularidad	# Comunidades
Red 1: 8M (día central)	0.859	221
Red 2: DiaDeLaMujer	0.780	475
Red 3: Marcha 8M	0.744	21
Red 4: Acoso	0.555	157
Red 5: Femicidio	0.833	69

Tabla 31. Modularidad y Comunidades por Red

Las principales Comunidades que se identificaron en dichas redes fueron las siguientes:

1. Política
2. Periodismo
3. Gobierno
4. Organizaciones de Derechos Humanos y Feministas
5. Organizaciones Deportivas / Culturales y Educativas
6. Otros

La participación de cada una de estas Comunidades por red se visualiza en la figura 31:

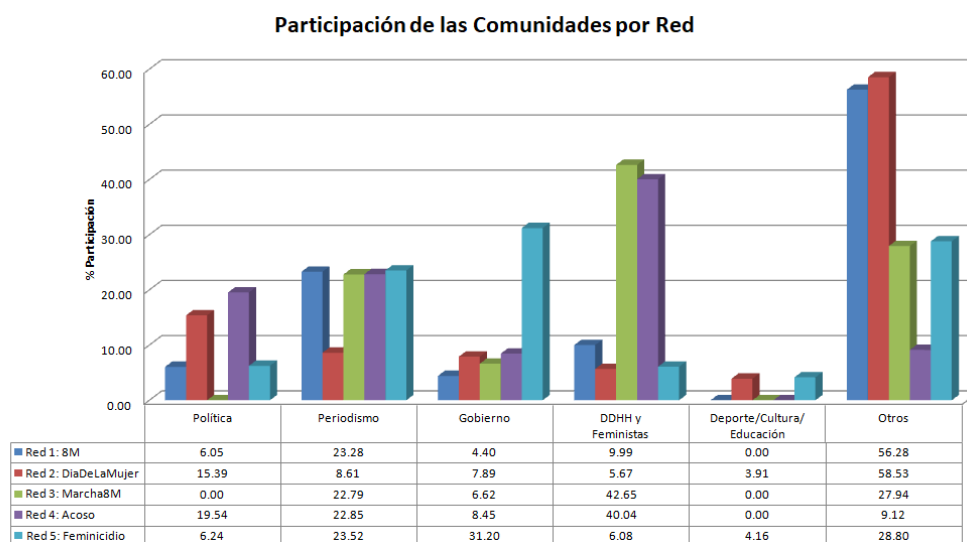


Figura 31. Tabla de participación de las Comunidades por Red

A partir de la tabla de la figura 31, se puede concluir lo siguiente:

- Si bien las redes 8M y Día de la Mujer tienen la misma connotación, se destaca que a nivel Político los comentarios en Twitter se promovieron más a través del hashtag #DíadelMujer en lugar que con el hashtag #8M, mientras que a nivel de la comunidad Periodismo sucedió lo contrario. Adicionalmente, se destaca que en la red Día de la Mujer, siendo más grande, (poco más de tres veces la cantidad de nodos que la red 8M), se puede encontrar mucho mayor nivel de relaciones entre sus participantes.
- La red de Acoso es la que menos porcentaje de nodos mantiene en la Comunidad denominada "Otros". Esto significa, que la mayoría de nodos están mucho mejor agrupados que en las otras cinco comunidades.
- La red de Femicidio es a la que mayor atención le presta la Comunidad de Gobierno y la de Periodismo.
- La Comunidad **Política** participó en mayor medida en la red #2 Día de la Mujer y en la red #4 Acoso. La Marcha8M (red #3) no tuvo acogida por la comunidad integrada por políticos, pero cabe señalar que algunos

integrantes de esta comunidad participaron en otras redes pero bajo el rol de activistas/feministas (es el caso de las Congresistas Indira Huilca y Marisa Glave).

- La Comunidad **Periodismo** ha participado generalmente de forma constante en las cinco redes, dándole mayor importancia y por tanto cobertura, al hashtag #8M (red #1) que al #DiadeLaMujer (red #2). El protagonismo del Periodismo en todas las redes es permanente.
- La Comunidad **Gobierno** mantiene presencia dentro de las cinco redes, pero se destaca en la red de Femicidio (red #5), que es en donde más ha participado y podría ser este tema social el que más le preocupa. Llama la atención que el Ministerio de la Mujer y Poblaciones Vulnerables ([@MimpPeru](#)) no apareciera dentro de los 10 nodos más influyentes en alguna de las redes analizadas.
- La Comunidad de **DDHH y Feministas** (organizaciones, colectivos y particulares), tiene una presencia marcada durante la Marcha 8M (red #3) e interactúa mucho con el tema de ACOSO hacia la mujer (red #4).
- La Comunidad **Deporte/Cultura/Educación** tiene escasa participación en las redes, sobresaliendo únicamente dentro de la red del Día de la Mujer (red #2) y dentro de la red de Femicidio (red #5).
- La Comunidad **Otros** reúne a todos aquellos nodos (usuarios de Twitter), que conforman comunidades muy pequeñas (menores al 3% del total de usuarios de la red) y por ende, no pueden graficarse de forma grupal.

5.3 Análisis Textual

Aprovechando el contenido textual de los retweets obtenidos en cada una de las redes, se ha identificado las palabras más relevantes por su frecuencia de aparición en el texto.



Figura 32. Word Cloud de la Red de Femicidio

La figura 32 muestra la nube de palabras obtenidas a partir de los retweets del usuario @saldelsol en la red de Femicidio.

Las 10 palabras más relevantes citadas en el texto son:

<u>Palabras</u>	<u>Frecuencia</u>
dia mujer	1.0
femicidio igualdad	0.9835164835164835
mujer dia	0.9615384615384616
acoso semana	0.9615384615384616
recuerda dia	0.9615384615384616
dia recuerda	0.9615384615384616
dia acoso	0.9615384615384616
semana femicidio	0.9615384615384616
igualdad dia	0.34065934065934067
femicidio mas	0.2087912087912088

Así mismo, para la red de Acoso, se visualiza la respectiva nube de palabras en la figura 33, pero esta no se filtró por un usuario relevante en particular.

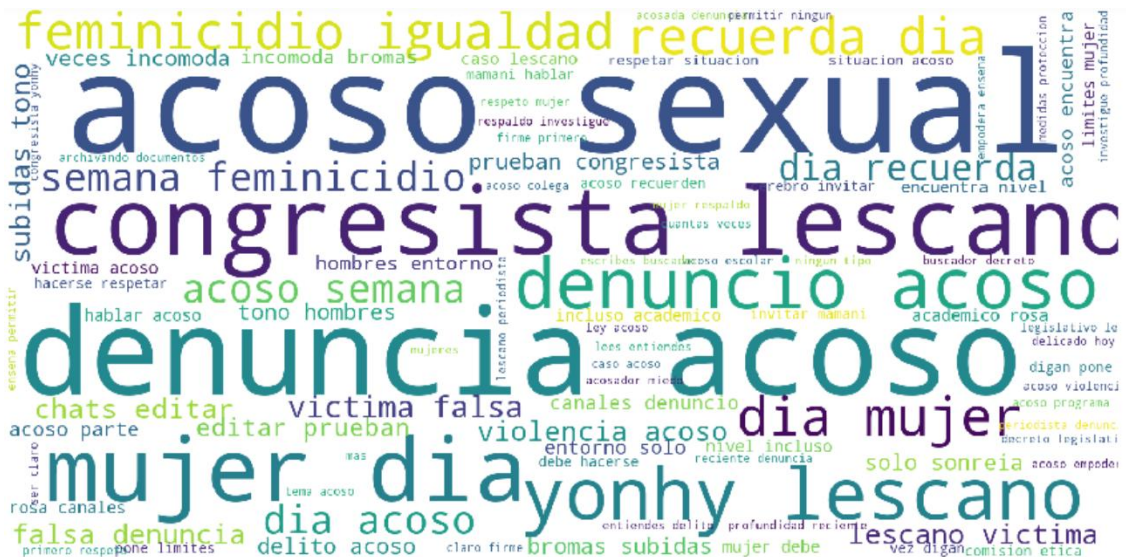


Figura 33. Word Cloud de la Red de Acoso

Y las 10 palabras más relevantes citadas en el texto son:

<u>Palabras</u>	<u>Frecuencia</u>
acoso sexual	1.0
denuncia acoso	0.7058823529411765
congresista lescano	0.48823529411764705
mujer dia	0.4264705882352941
yonyhy lescano	0.4176470588235294
denuncio acoso	0.40294117647058825
dia mujer	0.39705882352941174
feminicidio igualdad	0.38235294117647056
acoso semana	0.38235294117647056
recuerda dia	0.38235294117647056

Como principal observación en esta nube de palabras, destaca en tercer lugar las palabras “congresista lescano” y en quinto lugar “yonyhy lescano”. El congresista Yonhy Lescano, como se mencionó dentro de la red de Acoso, debe su aparición en esta red como actor relevante, debido a que previa a la fecha del 8 de marzo, específicamente el 4 de marzo último, fue denunciado por el delito de acoso sexual por una periodista radial. Debido a esta denuncia, fue separado de su cargo de Congresista de la Republica por un período de 120 días.

6. Conclusiones

El presente trabajo de fin de Maestría, ha logrado como objetivo principal el diseñar, modelar y analizar los datos de la red social Twitter a partir de los tweets recogidos entre el 4 al 12 de Marzo del 2019 (en el marco de celebración del Día Internacional de la Mujer), y que filtrados dentro del territorio peruano, dieron lugar a las siguientes cinco redes basadas en retweets y relacionadas al tema de la mujer peruana frente al acoso y el Femicidio:

- Red 8M
- Red Día de la Mujer
- Red Marcha8M
- Red de Acoso
- Red de Femicidio

Durante la carga de los tweets se trabajó con la librería Ujson en Python 3.7 (en lugar de la tradicional Json), y con ello, se logró optimizar el tiempo de procesamiento de los datos. Es importante mencionar que durante el proceso de filtrado de tweets dentro del territorio peruano, al no tener la mayoría de tweets el dato de geo localización registrado, puede que se haya omitido a algún actor relevante.

Por otro lado, el obtener la visualización de las redes fue inicialmente desordenado dentro de la plataforma de código abierto Gephi, y casi todo el proceso estuvo lleno de prueba y error. Al no existir la opción “deshacer” o ‘undo” (en inglés), para la mayoría de las acciones en la visualización de las redes, si se lograba una buena vista específica de los datos, se tuvo que exportar dicha vista (a un archivo de imagen por lo general), ya que con sólo presionar un botón se podía cambiar el resultado visual, inclusive si se comenzaba de nuevo con el mismo archivo de datos, era posible recrear ciertas vistas, pero esta tarea demandaba mucho tiempo.

Superada la fase de Minería de Datos y una vez obtenidos los modelos y las visualizaciones definitivas de las cinco redes propuestas (8M, Día de la Mujer, Marcha8M, Acoso y Femicidio), se calcularon los principales estadísticos, y a partir de las principales observaciones se obtuvieron las siguientes conclusiones:

1. La red 8M, fue analizada en tres periodos de tiempo, de los cuales la fecha central (8 de marzo el 2019), concentró a la gran mayoría de actores (nodos) y enlaces.
2. Las redes formadas por los hashtag 8M y Día de la Mujer, si bien trataron la celebración del Día internacional de la Mujer, en Perú, la red Día de la Mujer concentró poco más del triple de nodos y relaciones que la red 8M. Esta situación puede deberse a la activa participación de la comunidad Periodismo dentro de la Red Día de la Mujer, red que tuvo mayor cobertura periodística, mucho más que la Red 8M. Esta situación conlleva a pensar

que el papel del Periodismo es muy importante en la conformación de las redes sociales.

3. A nivel político, Keiko Fujimori ([@KeikoFujimori](#)), lideresa del Partido Político Fuerza Popular, y actualmente en prisión preventiva por lavado de activos entregados por la empresa constructora Odebrecht en 2011, demuestra que la privación de la libertad no impide una presencia activa en las redes sociales (dentro de la red Día de la Mujer). Incluso, se presenta como un personaje de igual o mayor influencia que el propio Primer Ministro peruano Salvador del Solar en dicha red.
4. La red Marcha8M congregó en su mayoría a las Organizaciones de Derechos Humanos y Feministas, y si bien contó con buena cobertura Periodística, esta no tuvo mayor relevancia a nivel Político y de Gobierno.
5. La red de Acoso fue la red que mejor agrupada se encontró en relación al resto de redes, esto quiere decir, que las comunidades estuvieron mejor distribuidas, con un bajo porcentaje de participación en la Comunidad Otros (9.12%).
6. La red de Femicidio es en la que el Gobierno tuvo mayor participación a través del nodo perteneciente al Primer Ministro de Perú, Salvador del Solar ([@saldelsol](#)).
7. Las Comunidades, que fueran detectadas siguiendo el algoritmo de clustering llamado "Modularity", e incluido en Gephi como Modularity Central, son las siguientes:
 - Política
 - Periodismo
 - Gobierno
 - Organizaciones de Derechos Humanos y Feministas
 - Organizaciones Deportivas / Culturales y Educativas
 - Otros
8. Por cada red, se identificaron los nodos más influyentes o importantes utilizando el algoritmo de centralidad de vector propio (Eigenvector Centrality), el cual mide cómo está conectada una entidad y la posible influencia directa que puede tener en la mayoría de entidades activas en la red.
9. El movimiento feminista #NiUnaMenos que se mencionara en el apartado 2.7 dentro del marco teórico, no aparece como relevante en ninguna de las redes analizadas, en su lugar, sí lo hace el movimiento NUMSTLosOlivos ([@NUMSTLosOlivos](#) NI UNA MENOS SOMOS TODAS - Los Olivos perteneciente al distrito limeño de Los Olivos), dentro de la red 8M.

En cuanto a la planificación y la metodología seguida para el desarrollo del trabajo, se destaca que las dos etapas más laboriosas, en las que se consumió

la mayor cantidad de tiempo, fue primero la etapa de captura de datos desde Twitter (Minería I) y la adecuación de estos al formato necesario para su posterior análisis (Minería II), y en segundo lugar, la etapa de modelado de las redes, que conllevó también un tiempo considerable para las observaciones y análisis.

Como líneas de desarrollo futuro se propone:

- Modelar la red de Feminismo/Feministas en Perú. Si bien en el presente trabajo se recogieron tweets con dichas palabras claves (feminismo y feministas), se desestimó el desarrollo de sus respectivas redes, porque en las cinco redes diseñadas se pudieron identificar a las principales organizaciones/movimientos y colectivos feministas peruanos.
- Realizar un análisis de sentimientos en base al texto de los tweets correspondientes a cada comunidad. Este análisis, puede servir a las entidades de gobierno para medir el sentimiento positivo, negativo o neutral de las personas/organizaciones (nodos) a partir de los tweets recogidos por ejemplo con las palabras clave Acoso y Femicidio.
- Hacer el seguimiento de la participación del gobierno, a través de las entidades gubernamentales dentro de futuras redes de Acoso y Femicidio a partir del 2020. La identificación de las palabras más relevantes citadas en el texto, nos dirán, por ejemplo, si las entidades de gobierno están presentes o no con programas de defensa a la mujer y poblaciones vulnerables. Mientras el sentimiento de la mayoría de nodos sea negativo, se entiende que no existe ningún cambio o lucha a favor de los derechos de las mujeres peruanas ante el acoso y feminicidio.

Al finalizar el presente TFM, se registran al día de hoy 67 casos de feminicidios entre enero y mayo del 2019 (recordemos que al inicio del trabajo, en el apartado 1.1. se contabilizaban sólo 22 casos de feminicidios).

El problema del acoso y feminicidios en Perú estará vigente en las redes sociales mientras no se encuentre una solución, y el Estado Peruano podría utilizar las redes sociales como un termómetro para medir el sentimiento de la población que participa preocupado por este problema, así como también como una herramienta para difundir sus estrategias de solución de una forma coordinada, entre todas las entidades gubernamentales involucradas en la defensa de los de los derechos de la mujer y las poblaciones vulnerables del Perú.

7. Glosario

API: Application Programming Interface o interfases de programación de aplicaciones. Un API permite la comunicación entre diferentes componentes de software.

Back-end: El desarrollo de back-end se refiere al lado del servidor de una aplicación y todo lo que se comunica entre la base de datos y el navegador.

Bot: Un bot (aféresis de robot) es un programa informático que efectúa automáticamente tareas repetitivas a través de Internet.

Csv: Los archivos CSV (del inglés comma-separated values) son un tipo de documento en formato abierto sencillo para representar datos en forma de tabla, en las que las columnas se separan por comas y las filas por saltos de línea.

Dataframe: es la estructura de datos básica que proporciona la librería Pandas en Python. Está conformada por una colección ordenada de columnas con nombres y tipos, parecido a una tabla de base de datos, donde una sola fila representa un único caso (ejemplo) y las columnas representan atributos particulares (pudiendo ser de tipos de datos diferentes).

Emoji: es un término japonés para los ideogramas o caracteres usados en mensajes electrónicos y sitios web

Feminicidio: es el homicidio perpetrado hacia una mujer por motivo de violencia familiar, sexual o discriminación de género. El Feminicidio, definido como la muerte de mujeres a causa de la discriminación de género, constituye la manifestación más cruel que adopta la violencia contra la mujer y representa un grave problema Social.

Gephi: software open-source de análisis de redes y visualización escrito en Java.

Hashtag: es la palabra o conjunto de palabras o caracteres alfanuméricos que precedidos por el símbolo de numeral (#), se utiliza en Twitter y en otras redes sociales en Internet, para clasificar o agrupar el contenido de las publicaciones. Esta palabra proviene del inglés, que traducida al castellano, significa 'etiqueta'.

INEI: Instituto Nacional de Estadística e Informática en Perú.

Internet: el término se forma a partir de dos palabras en inglés Inter(national)net(work). De acuerdo a la definición de Internet de la Real Academia Española (RAE), se trata de una red informática compuesta por un sistema de redes de comunicación interconectadas por una familia de protocolos TCP/IP. Esto hace que las redes físicas formen una red lógica de alcance mundial. (Gestión, 2018)

JSON: JavaScript Object Notation o Notación de Objetos de JavaScript en castellano, es un formato ligero de intercambio de datos. JSON es un formato de texto que es completamente independiente del lenguaje pero utiliza convenciones que son ampliamente conocidos por los programadores de la familia de lenguajes C, incluyendo C, C++, C#, Java, JavaScript, Perl, Python, y muchos otros. Estas propiedades hacen que JSON sea un lenguaje ideal para el intercambio de datos.

MIDIS: Ministerio de Desarrollo e Inclusión Social en Perú.

MINITER: Ministerio del Interior en Perú.

MIMP: Ministerio de la Mujer y Poblaciones Vulnerables en Perú.

Pandas: es una librería en Python para el análisis de datos que cuenta con las estructuras de datos necesarios para limpiar los datos en bruto y que sean aptos para el análisis.

PEC: Pruebas de Evaluación Continua.

PNP: Policía Nacional del Perú.

Python: lenguaje de propósito general compuesto por herramientas / bibliotecas e ideal para el desarrollo web de back-end, el análisis de datos, la inteligencia artificial y la computación científica.

RT: Abreviación para ReTweet. Es la acción de compartir o difundir un tweet proveniente de otro usuario al que se sigue.

SNA: Social Network Analysis por sus siglas en inglés para Análisis de Redes Sociales, es una técnica clave en muchos campos como la sociología, la biología, comunicación, política que permite analizar relaciones entre elementos para determinar la importancia de los actores, cómo se han agrupado en subcomunidades, el flujo de la información, lo que es clave en la transmisión de la información, etc.

TFM: Trabajo de fin del Master.

UJSON: librería en Python llamada UltraJSON, que es un codificador y decodificador JSON ultrarrápido escrito en C pura.

UNMSM: Universidad Nacional Mayor de San Marcos, en Lima - Perú

URL: son las siglas en inglés de Uniform Resource Locator, que en español significa Localizador Uniforme de Recursos. Como tal, una URL es la dirección específica que se asigna a cada uno de los recursos disponibles en la red de Internet con la finalidad de que estos puedan ser localizados o identificados.

Web: vocablo de origen inglés que significa “red”, “telaraña” o “malla”. En el ámbito tecnológico, es utilizado para identificar una red informática y a la propia Internet.

Word Cloud: es una imagen compuesta de palabras utilizadas en un texto o tema en particular, en el que el tamaño de cada palabra indica su frecuencia o importancia.

WWW: El World Wide Web (WWW) es una red de “sitios” que pueden ser buscados y mostrados con un protocolo llamado HyperText Transfer Protocol (HTTP). [2] Se trata de un sistema de distribución de documentos de hipertexto o hipermedios interconectados y accesibles vía Internet. Cualquier usuario de Internet puede ver dichos archivos, sea texto, vídeo, foto u otros y navegar entre ellos. [5]

8. Bibliografía

- [1] INEI. (2018, Setiembre). *Perú: Estadísticas de Femicidio 2011-2018 Registros Administrativos - Síntesis Estadística*. Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/boletines/boletin_femicidio2018.pdf
- [2] KEMP, S. (2019, Enero 30). *DIGITAL 2019: GLOBAL INTERNET USE ACCELERATES*. Retrieved Marzo 18, 2019, from <https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates>
- [3] FIB, F. d. (2002). *Historia de internet*. Retrieved 2019, from <https://www.fib.upc.edu/retro-informatica/historia/internet.html>
- [4] Netcraft. (2019, Enero). *January 2019 Web Server Survey*. Retrieved from <https://news.netcraft.com/archives/2019/01/24/january-2019-web-server-survey.html>
- [5] Gestión. (2018, Julio 31). *¿Cuál es la historia de Internet?* Retrieved from <https://gestion.pe/tecnologia/historia-internet-240094?ref=gesr>
- [6] Rouse, M. (2018, Noviembre). *Big data, Maximizing and managing big data with SOA middleware*. Retrieved Marzo 18, 2019, from <https://searchdatamanagement.techtarget.com/definition/big-data>
- [7] Laney, D. (2001, February 6). "3D Data Management: Controlling Data Volume, Velocity, and Variety", *Gartner*. Retrieved from File No.949: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity- and-Variety.pdf>
- [8] Patgiri, R. &. (2016). *Big Data: The V's of the Game Changer Paradigm*. Retrieved from 10.1109/HPCC-SmartCity-DSS.2016.0014.: https://www.researchgate.net/publication/311642627_Big_Data_The_V's_of_the_Game_Changer_Paradigm
- [9] Wikipedia. (2019, Marzo 26). *Chasqui*. Retrieved Marzo 15, 2019, from <https://es.wikipedia.org/wiki/Chasqui>
- [10] McFadden, C. (2018, Octubre 16). *A Chronological History of Social Media*. Retrieved 2019, from <https://interestingengineering.com/a-chronological-history-of-social-media>
- [11] ESRC, E. a. (2019). *What is Twitter and why should you use it?* Retrieved Marzo 14, 2019, from <https://esrc.ukri.org/research/impact-toolkit/social-media/twitter/what-is-twitter/>
- [12] Sturk, C. (2018, Noviembre 12). *A Look at the Many Different Uses of Twitter*. Retrieved Marzo 14, 2019, from <https://www.lifewire.com/what-is-twitter-for-3288888>

- [13] Jorge, M. (2018, Abril 30). *Parece que Twitter también vendió el acceso a sus datos a Cambridge Analytica*. Retrieved Marzo 15, 2019, from <https://es.gizmodo.com/parece-que-twitter-tambien-vendio-el-acceso-a-datos-a-1825645545>
- [14] Kapoor, A. (2018, Octubre 2). *Análisis de redes sociales en Python*. Retrieved Marzo 16, 2019, from <https://www.datacamp.com/community/tutorials/social-network-analysis-python>
- [15] Clauset, A. (2015). *A Short Introduction to Networks*. Retrieved Marzo 19, 2019, from University of Colorado Boulder: http://tuvalu.santafe.edu/~aaronc/slides/Clauset_2015_SFI_Networks_ShortCourse.pdf
- [16] Girvan M, N. M. (2002). Community structure in social and biological networks. Proceedings of the national academy of sciences. In N. M. Girvan M, *Community structure in social and biological networks. Proceedings of the national academy of sciences* (pp. 7821-7826). USA: Proc. Natl Acad. Sci.
- [17] Importancia. (2019). *Importancia del Feminismo*. Retrieved Marzo 19, 2019, from <https://www.importancia.org/feminismo.php>
- [18] Sarah J. Jackson, S. B. (2016, Septiembre 8). *Digital Standpoints: Debating Gendered Violence and Racial Exclusions in the Feminist Counterpublic*. Retrieved Marzo 23, 2019, from <https://doi.org/10.1177/0196859916667731>
- [19] Barker-Plummer, D. &.p. (2017). Social Movements and Media. In *Twitter as a Feminist Resource: #YesAllWomen, Digital Platforms, and Discursive Social Change* (pp. 91-118). 10.1108/S2050-206020170000014010.
- [20] Cosmo, C. (2019). *Glosario de términos para Twitter*. Retrieved Marzo 31, 2019, from Cosmo Ciudadano, Comité de observación, investigación y desarrollo para Internet: <http://www.cosmociudadano.mx/glosario-de-terminos-para-twitter/>
- [21] Alcaraz, M. F. (2017, Septiembre 4). [SPECIAL EDITION] #NIUNAMENOS: *POLITICISING THE USE OF TECHNOLOGIES*. Retrieved Marzo 27, 2019, from <https://www.genderit.org/feminist-talk/special-edition-niunamenos-politicising-use-technologies>
- [22] Peru, D. (2019, 05 05). *Defensoria del Pueblo*. Retrieved 05 05, 2019, from Defensoria del Pueblo: <https://www.defensoria.gob.pe/quienes-somos/>

9. Anexos

Anexo 1: Planificación del Trabajo – Diagrama de Gantt

