

Dirección de Desarrollo Digital

Unidad de Científicos
de Datos



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación



Análisis de Tendencias sobre el uso del SIPO 2.0

INFORME FINAL

Dependencias y entidades involucradas	Departamento Nacional de Planeación <ul style="list-style-type: none">• Dirección de Desarrollo Digital - Unidad de Científicos de Datos• DSEPP – Grupo de Seguimiento de Paz y Estabilización; Dirección General / Grupo de proyectos Especiales – Equipos de Paz y Víctimas
Sector	Justicia
Tecnologías utilizadas	API tweepy, Python
Fuentes de datos	<ul style="list-style-type: none">• Lista de términos relacionados al sistema SIPO – DSEPP• Portales de Noticias y secciones – DSEPP• Tweets diarios desde 24/02/2021 – 21/05/2021• Noticias diarias desde 08/03/2021 – 19/05/2021

Contenido

1. Presentación	2
2. Objetivos del proyecto	2
3. Metodología	3
4. Resultados	5
5. Conclusiones y recomendaciones	19
6. Socialización	19
Contacto	20



1. Presentación

La principal motivación del proyecto consiste en la generación de insumos de datos para el Grupo de Seguimiento de Paz con el fin de que estos sirvan para la generación de un plan de socialización y capacitación SIIPO 2.0 (Sistema integrado de información para el postconflicto), de tal manera que se haga eficiente el uso de información disponible sobre el posconflicto en Colombia, esta se realiza a partir de la identificación de tendencias de contenido relacionado a SIIPO en portales de noticias y posts dentro de la red social Twitter a través de técnicas de Procesamiento de Lenguaje Natural y Web scrapping.

Se encontró que los tweets que contienen términos específicos como “Implementación al Acuerdo”, “SIPPO” etc., suelen tener contenido que incita al rechazo de los Acuerdos de Paz realizados en Colombia y por ende se observa una recepción altamente negativa hacia los Acuerdos de Paz a nivel general, donde los principales influenciadores de contenido relacionados a estos son individuos que participan en la Política Colombia, activistas políticos y especialistas en opinión política. Sin embargo, esto no necesariamente implica que estos influenciadores sean los generadores de estas opiniones que convergen, al considerar su nivel de influencia explícitamente a través de la cantidad de Retweets acumulados que tienen en sus posts. En cuanto a las noticias, se encontró que los temas de conversación para las columnas de opinión, política y académica se centran en temas de coyuntura colombiana y no en opiniones relacionadas al sistema SIIPO, en este orden de ideas tanto para tweets como para los artículos de portales de noticias, las menciones de SIIPO son muy inferiores a las menciones de la frase “Acuerdo de Paz”, además de cercanas a cero a través del tiempo, por lo que la generación de contenido de opinión y análisis sobre el Sistema de Implementación para el Posconflicto es muy pequeño y los términos de búsqueda conducen a opiniones sobre los Acuerdos de Paz y en una pequeña cuantía hacia el sistema SIIPO.

The main motivation of the project is the generation of data inputs for the Peace Monitoring Group so that these inputs can be used to generate a SIIPO 2.0 socialization and training plan in order to make efficient use of available information on post-conflict in Colombia, thus identifying trends in content related to SIIPO in news portals and posts within the social network Twitter through Natural Language Processing and Webscraping techniques.

It was found that tweets containing specific search terms such as "Implementation to the Agreement", "SIPPO" etc., have a tendency towards content that incites the rejection of the Peace Agreements made in Colombia and therefore a highly negative reception towards the Peace Agreements is observed at a general level, where the main influencers of content related to these are individuals involved in Colombia Politics, political activists and specialists in political opinion, however this does not necessarily imply that these influencers are the generators of these opinions that converge, when considering their level of influence explicitly through the amount of accumulated Retweets they have in their posts. And for news, it was found that the topics of conversation for opinion, political and academic columns focus on Colombian conjuncture topics and not on opinions related to the SIIPO system, in this order of ideas for both tweets and news portal articles, Mentions of SIIPO are much lower than mentions of the word Peace Agreement and close to zero over time, so the generation of opinion and analysis content on the Post-Conflict Implementation System is very small and search terms lead to opinions on the Peace Agreements and in a small amount on SIIPO.

2. Objetivos del proyecto

2.1. General

Realizar un análisis de tendencias sobre las menciones e interacciones con el Sistema Integrado de Información para el Posconflicto - SIIPO que brinden insumos al Grupo de Seguimiento a Paz y Estabilización para generar un plan de socialización y capacitación del SIIPO 2.0 que permita potenciar el uso de la información disponible.



2.2. Específicos

1. Identificar los actores/influenciadores que generan contenido relacionado con los avances en la implementación del Acuerdo Final de Paz hacen uso del SIPO en sus menciones y publicaciones.
2. Conocer las tendencias relacionadas con el uso del SIPO en distintas fuentes de información, así como los contextos en los que se hace mención al uso del sistema y la información disponible.

3. Metodología

La metodología puede dividirse en varias fases las cuáles fueron necesarias para obtener todos los insumos que permitieran generar conclusiones, en base a la naturaleza exploratoria del análisis los resultados permiten obtener una aproximación inicial respecto a las principales tendencias de los tweets asociados al sistema SIPO 2.0 y los artículos de noticias con las mismas características. Inicialmente se realizó un proceso de Web scrapping, el cual consiste en extraer información de sitios web, para su posterior almacenamiento en bases de datos estructuradas o no estructuradas, y finalmente el análisis de esta. Este proceso fue aplicado para los artículos de noticias y para la descarga de tweets se utilizó la API de Twitter Tweepy. Los términos de búsqueda tanto para noticias y tweets son los siguientes.

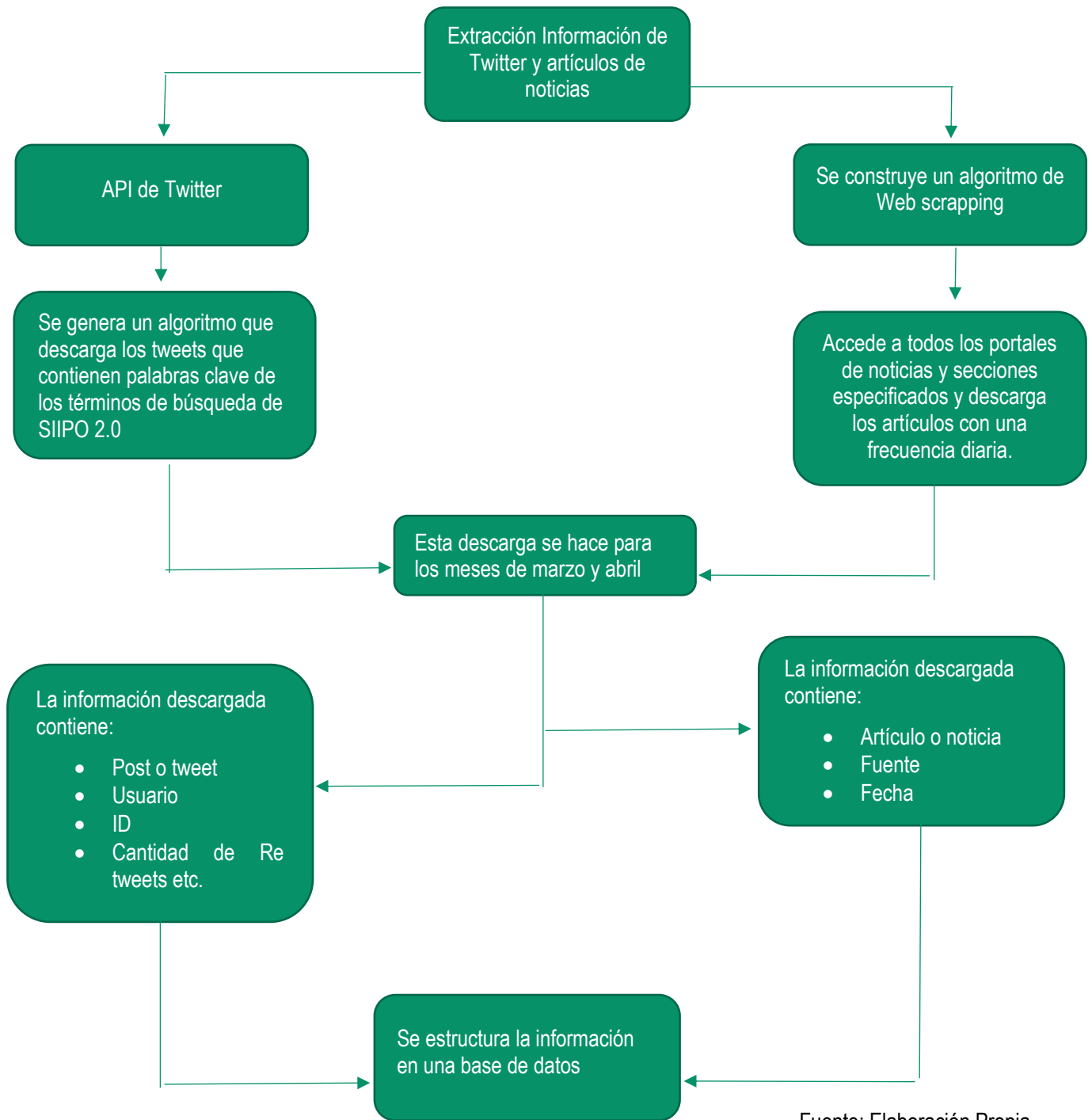
1. Lista de términos relacionados al sistema SIPO para realizar la descarga de tweets:
 - SIPO
 - SIPO
 - Sistema Integrado de Información para el Posconflicto
 - Seguimiento a la implementación
 - Seguimiento al Acuerdo de Paz
 - Avance Implementación Acuerdo de Paz
 - Seguimiento al Acuerdo Final
 - Cifras de implementación del Acuerdo Final
 - Informes de seguimiento al Acuerdo de Paz

Por su parte, la lista de portales de noticias es la siguiente:

2. Portales de noticias y secciones de las cuales se descargará la información para su análisis:
 - El Tiempo.com / Política
 - Eltiempo.com / Opinión
 - Espectador.com / Colombia2020
 - Espectador.com / Opinión
 - La silla vacía / Silla Nacional
 - La silla vacía / Silla Académica

A partir de estos términos de búsqueda relacionados a SIPO y los portales de noticias, se genera el proceso de descarga tanto de artículos de noticias como de los tweets. En la Figura 1 se describe el proceso de forma más detallada:

Figura 1: Flujo de Extracción de Información para tweets y artículos de noticias



3.1. Generación de gráficos de frecuencia y nubes de palabras

Posterior a la recolección, almacenamiento y consolidación de las bases de datos, se genera un proceso de limpieza de datos. Este consta de los siguientes pasos:

- Se guardan en una lista todos los usuarios de Twitter cuyos tweets se encuentran en la base de datos.
- Se guardan los términos de búsqueda de SIPO en una lista



- Se guardan los nombres de los portales de noticias utilizados en una lista
- Se descargan diccionarios de stopwords, que contienen los nombres, apellidos del español, los nombres de ciudades y municipios de Colombia y otras palabras que por sí mismas no tienen ningún significado, por ejemplo, conectores lógicos como “sí”, “cuándo”, “solo sí” etc. Este conjunto de palabras consolidadas en una lista única, son eliminadas de los textos, añadiendo las listas de usuarios de Twitter, los términos de búsqueda a su vez como los signos de puntuación, mayúsculas, expresiones especiales (@, //, etc.), igual que todas aquellas expresiones de menos de 4 caracteres. Este proceso busca eliminar el ruido y palabras que no aportan significado para los textos.

Para los gráficos de frecuencias, se utiliza la librería Contexto, elaborada en Python por la UCD para el análisis de texto, donde generan las gráficas de frecuencias de palabras para unigramas y bigramas, esto quiere decir que muestra de forma gráfica la frecuencia de palabras únicas (como casa, hogar, exclusión etc.) y compuestas (como obra pública, salud pública etc.) para los tweets y los artículos de noticias. Mostrar las palabras más frecuentes de los textos permiten dar una idea de los principales temas y opiniones que existen respecto a los tweets y artículos de noticias asociados a los términos de búsqueda de SIIPO.

3.2. Generación de gráficos de palabras en el tiempo y análisis complementarios

El proceso anterior permite conocer cuáles son los principales términos más frecuentes en noticias y tweets. A continuación, se realizan los siguientes pasos para generar insumos para el análisis:

- Se realiza un conteo en cada uno de los textos de las palabras únicas y compuestas más frecuentes y se adicionan en la base de datos. Esto se realiza tanto para los tweets como para los artículos de noticias, este conteo se realiza posterior a la limpieza de los textos, para asegurar la mayor fiabilidad posible de reconocimiento de estas palabras más frecuentes.
- Se agrupa la base de datos en distintas temporalidades que pueden ser interesantes igual que por término de búsqueda para el caso de los tweets y por fuente para los artículos de noticias. Las series de tweets se analizan de forma diaria y las series de noticias se analizan de forma semanal, de tal manera se pueda observar la evolución de las palabras más frecuentes para tweets y noticias en el tiempo.
- Se agrupa la base de datos de Twitter por usuario y número de Retweets acumulados como un proxy de los actores más influyentes que generan contenido asociado a SIIPO en Twitter durante los dos últimos meses, considerando que entre mayor sean la cantidad de Retweets mayor es la dispersión de información en el público.
- Se genera una reducción de dimensionalidad a través del algoritmo t-SNE, con el fin de entender de forma global las frecuencias de las palabras más recurrentes en los tweets y artículos de noticias en un espacio bidimensional y de esta manera ver la relación entre las mismas.
- Se genera una matriz y un gráfico de Coocurrencias que permite ver la relación que existe entre las diferentes palabras más frecuentes (Unigramas y Bigramas) y si esta relación es débil o fuerte, es decir si los términos relacionados son mutuamente dependientes lo que indica que los textos se encuentran comúnmente de forma conjunta.

4. Resultados

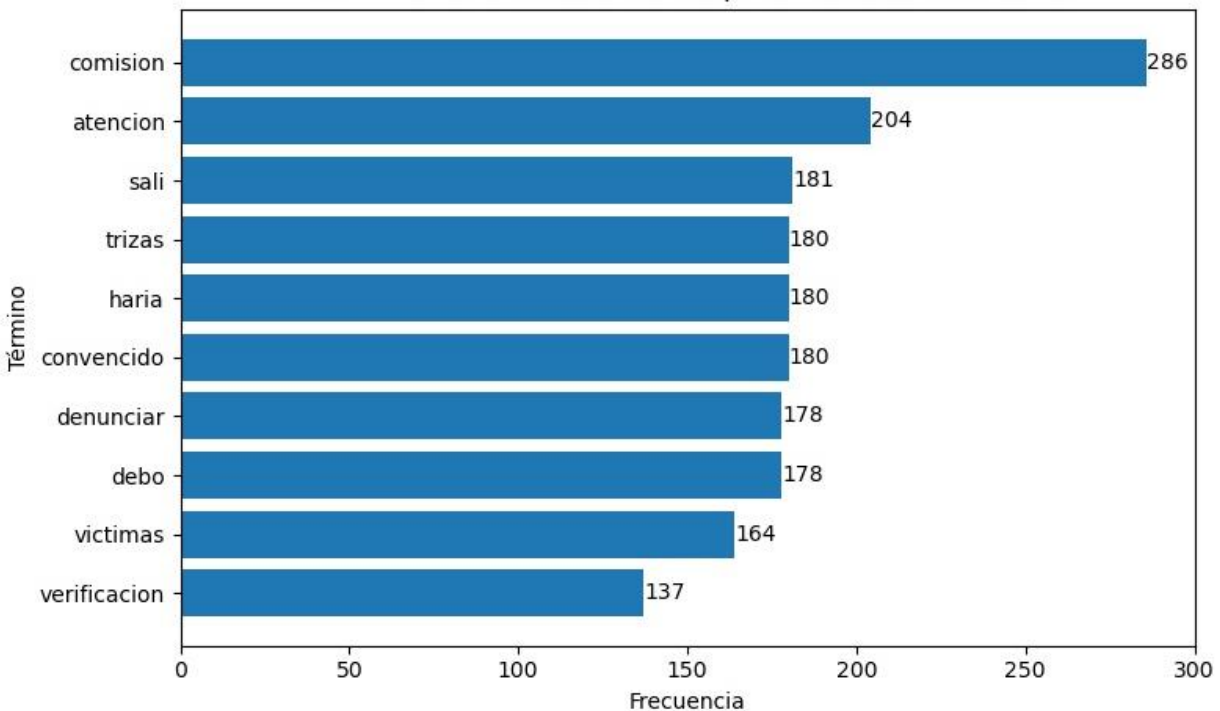
A través del desarrollo metodológico descrito en la Sección 3 se obtuvieron los resultados que se presentan a continuación. Toda retroalimentación desde un punto de vista experto o de usuario por parte de la DSEPP es bienvenida. Este insumo será de gran ayuda para mejorar la calidad y utilidad de los resultados obtenidos, de manera que agreguen mayor valor.



Inicialmente, a través de la exploración realizada en los datos se determinan que las palabras más frecuentes de los artículos de noticias en las columnas especificadas y los tweets con la filtración de contenido asociada a las palabras clave de SIPO, se tienen los siguientes resultados:

4.1 Análisis de Tweets

Gráfico 1: Términos únicos más frecuentes de los Tweets Asociados a SIPO



Fuente: Elaboración Propia

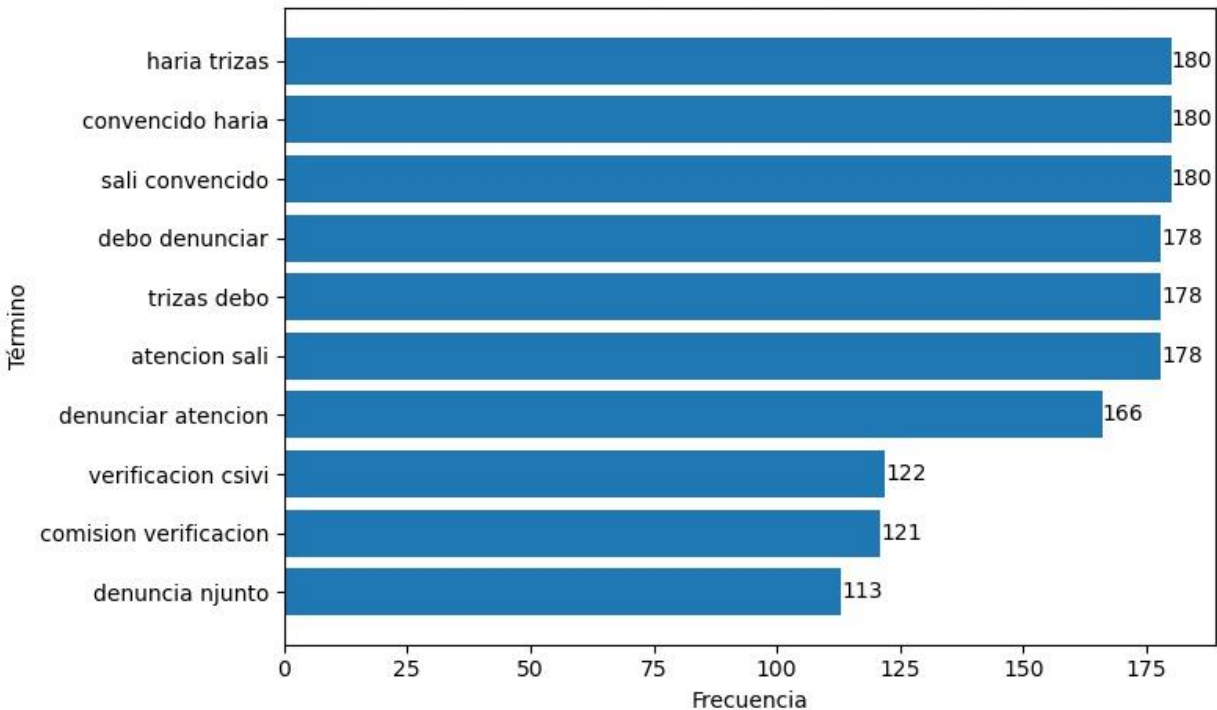
A partir del Gráfico 1, que consiste en las principales palabras más frecuentes de los tweets asociados a SIPO, podemos observar que palabras como “salir”, “trizas”, “hacer”, “convencido” son asociadas a sentimientos negativos que posteriormente en el gráfico de bigramas de Twitter se podrá entender mejor su connotación. Se podría en primera instancia pensar que estas palabras pueden estar asociadas con el Sistema Integrado de Información para el Posconflicto SIPO, sin embargo existe una ambigüedad ya que puede referirse a los acuerdos de paz firmados en Colombia en el año 2014 y no directamente al sistema SIPO, por lo que esta hipótesis la exploraremos más adelante para poder entender de forma precisa el contexto y el objeto al cuál se dirige este contenido.

Por otra parte, entonces parece que el tema más frecuente que se habla a partir de la palabra de mayor frecuencia de aparición se refiere a la “Comisión de Verificación”, que puede asociarse a la participación de la ONU en el posconflicto de Colombia a través de la misión de verificación. La cuál es una misión de carácter internacional, que funciona como órgano de control independiente para verificar que se cumplan las condiciones definidas en el Acuerdo de Paz tanto para el Estado y las víctimas del postconflicto, así como para las FARC. Por ejemplo, dentro de las labores de esta comisión está la de verificar el cumplimiento de la reincorporación de las FARC a la vida civil, económica y política, dando garantías para el nuevo partido político FARC-EP, por otro lado, la garantía de seguridad y protección para todos los involucrados dentro del acuerdo de paz y la implementación de este. De esta manera este organismo, busca



afianzar la confianza entre las partes y transparencia internacional frente a todo aquello que ocurra alrededor de los Acuerdos de Paz.

Gráfico 2: Términos compuestos más frecuentes de los Tweets Asociados a SIIPO



Fuente: Elaboración Propia

A partir del Gráfico 2 se puede identificar que la palabra más frecuente es “haría trizas”, seguido de palabras que tienen una connotación similar, por otra parte, la palabra “debo denunciar” y “atención salir” pueden ser asociadas al deseo de la manifestación de protesta por parte de los usuarios de Twitter. Sin embargo, es necesario aclarar en base a la hipótesis presentada anteriormente, si la concentración de estos sentimientos está asociada al sistema Integrado de Información para el Posconflicto SIIPO o para los Acuerdos de Paz, de tal manera como se puede observar en el Gráfico 3.

Tenemos que el comportamiento de la frecuencia de la palabra SIIPO en los tweets se concentra en valores pequeños cercanos a cero y mantiene este comportamiento de forma constante a lo largo de Febrero hasta el mes de Mayo, por el contrario, el término “Acuerdo de Paz”, tiene un comportamiento similar al de la palabra SIIPO a lo largo del tiempo. Sin embargo, aproximadamente el 4 de Abril tiene un crecimiento exponencial, llegando a aparecer por encima de las 200 veces en los diferentes tweets. De esta manera para los tweets teniendo en cuenta el comportamiento de ambas variables a lo largo del tiempo, especialmente la tendencia de la palabra SIIPO a concentrarse en valores cercanos a cero, podemos inferir que las palabras “Hacer Trizas”, “Debo Denunciar”, “Comisión de Verificación” se convierten en el principal contenido y temas centrales de los tweets dirigidos hacia los Acuerdos de Paz de Colombia y no hacia el Sistema Integrado de Información SIIPO.

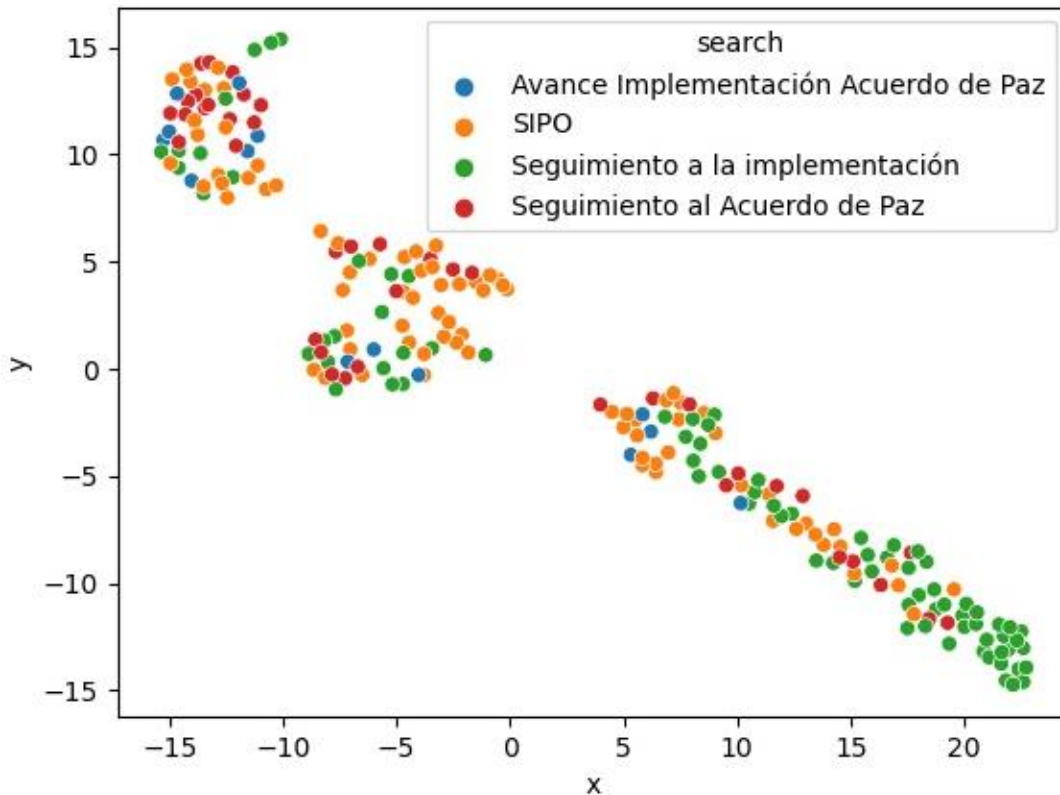


El Gráfico 4 nos muestra tres clústeres (grupos) principales donde existe una alta relación (una probabilidad del 90% de que las palabras estén relacionadas y por ende puedan aparecer de forma conjunta comúnmente en los tweets) entre palabras, intuitivamente los nodos o puntos en este caso representan palabras y las líneas o aristas representan la relación entre los diferentes nodos, el grosor de la arista implica que existe una alta relación entre dos palabras y por ende es probable que sean utilizadas comúnmente de forma conjunta.

En el primer clúster, las palabras “Salir Convencido”, “Debo Denunciar”, así como en el segundo clúster palabras como “Conflicto Víctimas”, “Armado Víctimas”, “Comisión Verificación” etc. Permiten observar intuitivamente aquellas palabras que se unen para generar unidades compuestas más complejas dentro del texto, de esta manera, a nivel general para los tweets las palabras compuestas comunes, se establecen en tres clústeres diferentes. Estas palabras cobran sentido dentro del marco en el cual estas se expresan (acuerdos de paz) que hablan en su mayoría hacia los acuerdos, mas no sobre SIPO.

Finalmente, en congruencia con los resultados que hemos observado anteriormente estas palabras que tienen una relación alta pueden asociarse a un contexto de conflicto, denuncias, víctimas por lo que implica que las palabras en estos tweets en base a los términos de búsqueda de SIPO se constituyen alrededor de contenido asociado al conflicto armado y concentrados en la discusión en la red social sobre los Acuerdos de Paz.

Gráfico 5: t-SNE para los tweets



Fuente: Elaboración Propia

A partir del Gráfico 5, vemos que todos los puntos en el espacio se agrupan de forma conjunta entre los distintos términos de búsqueda relacionados a SIPO, estos puntos en el espacio se refieren a cada uno de los valores de frecuencia de aparición en el tiempo para las distintas palabras únicas y compuestas en un espacio bidimensional,



esto permite ver qué tipo de relación existe entre las diferentes variables cuantitativas que pueden estar en escalas de medición distintas.

Este agrupamiento de las diferentes palabras por término de búsqueda probablemente se debe a que las palabras de alta frecuencia para el caso de Twitter se encuentran en general para los diferentes términos de búsqueda, es decir, que los tweets con los términos de búsqueda “Avance Implementación Acuerdo de Paz”, “Seguimiento a la Implementación”, “Seguimiento al Acuerdo de Paz” y “SIIPO” tienen de forma general las mismas palabras comunes y por ende se agrupan de esta manera. Estas palabras agrupadas, se pueden visualizar en el Gráfico 5 y 6 y representan entonces las palabras de los clústeres, anteriormente explicados. De esta manera el tamaño de las palabras se asocia con la frecuencia de aparición de la misma dentro de los tweets, es decir entre más grande sea su tamaño mayor es su frecuencia de aparición.



Gráfico 5 y 6: Nubes de palabras de palabras únicas y compuestas para los tweets

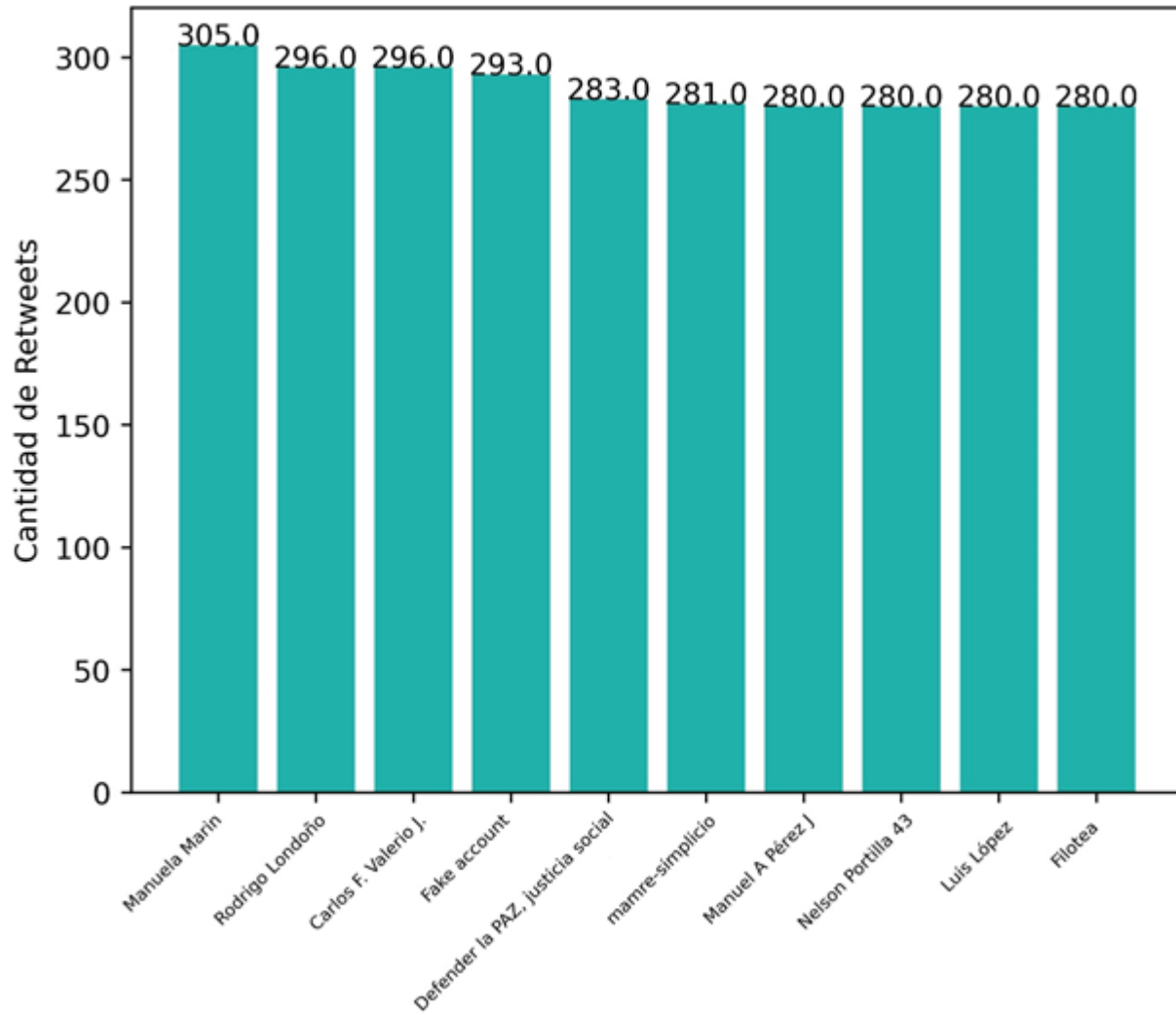


Fuente: Elaboración Propia

Finalmente, es importante determinar los agentes influenciadores que generan contenido relacionado con SIPO, para el análisis de influenciadores de tweets, se generó una variable que es una aproximación al nivel de difusión de información por cada usuario que posteo contenido relacionado a SIPO. En este caso nos referimos a la cantidad de Retweets acumulados por individuo a lo largo de abril y mayo para los posts generados con contenido relacionado a SIPO, como se puede observar a continuación:



Gráfico 7: Principales Influenciadores del sistema SIPO 2.0 en Twitter



Fuente: Elaboración Propia

A partir del Gráfico 7, tenemos los principales generadores de contenido relacionado al sistema de SIPO 2.0, Manuela Marin es la que mayor difusión ha generado a la fecha, su contenido a nivel agregado ha generado 305 Retweets, directamente es una persona que trabaja en el ámbito del activismo político, siendo integrante de la Dirección Distrital Comunes de Bogotá, su línea de trabajo está relacionada a la regulación y verificación del acuerdo de paz en la Habana.

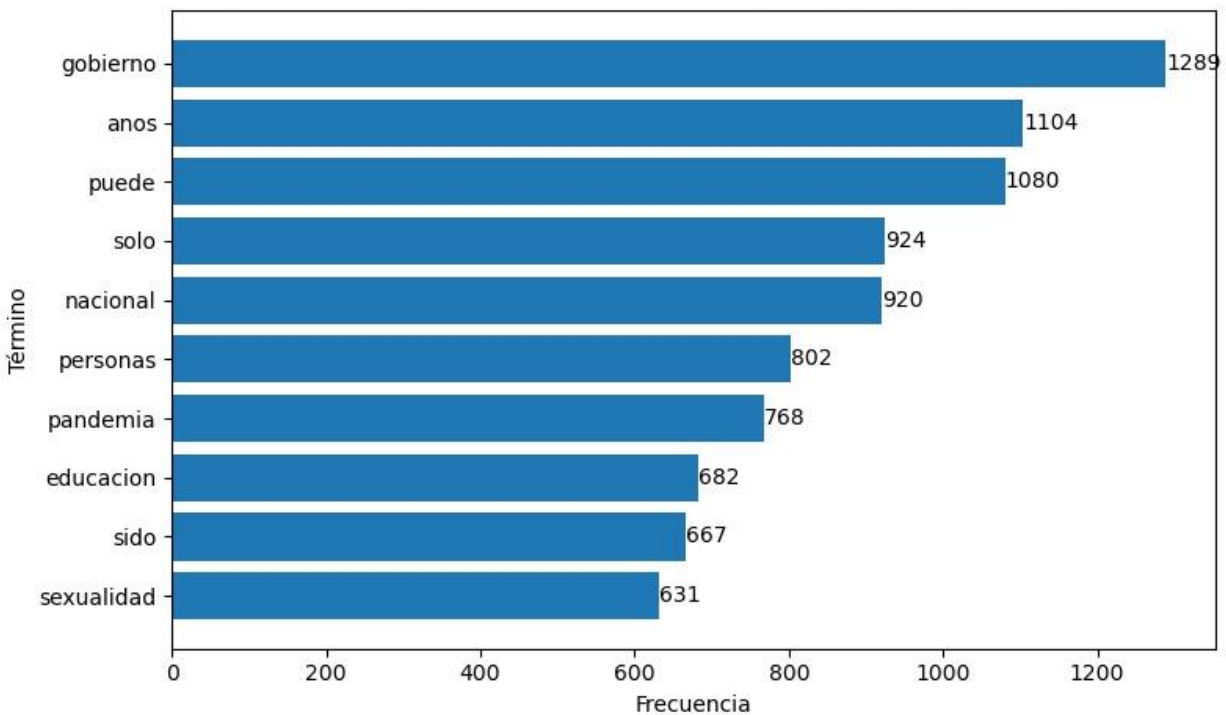
Por otra parte Rodrigo Londoño tienen una influencia de 296 Retweets acumulados en sus post, un ex comandante de las FARC que pasó a participar en la política colombiana y fue orientador de la delegación de paz que terminó pactando el acuerdo general para la terminación del conflicto, de esta manera los principales influenciadores son personas asociadas a la política y el activismo político en Colombia y a su vez individuos con afinidades hacia la participación de opinión en política como mamre-simplicio y Filotea. A pesar de ser los principales influenciadores identificados, esto no implica que el contenido que generen ellos esté en la misma línea de pensamiento que predomina en el contenido asociado a los términos de búsqueda de SIPO a nivel general.



Finalmente, para los tweets, podemos concluir que los temas relacionados al conflicto armado en Colombia que se discuten en la red social, están asociados a los Acuerdos de Paz al ser esta la palabra que mayores menciones tiene en los tweets a comparación de la palabra SIIPO, por otra parte, el contenido de los tweets basados en los términos de búsqueda de SIIPO sugiere un rechazo y un sentimiento negativo o de malestar, respecto a la forma en que están pactados los acuerdos de paz en Colombia.

4.2 Análisis de Artículos de noticias

Gráfico 8: Palabras únicas más frecuentes de Artículos de Noticias

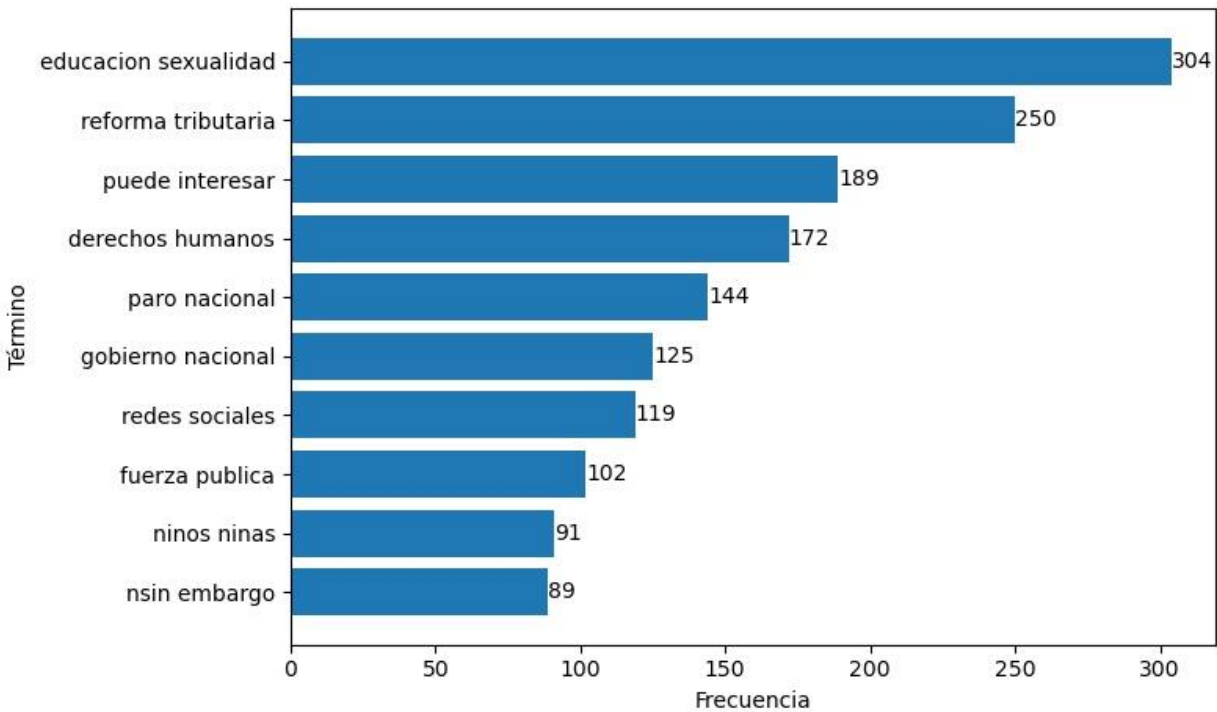


Fuente: Elaboración Propia

A partir del Gráfico 8, podemos observar que no existen para las palabras únicas términos asociados a SIIPO 2.0, a partir de la búsqueda en las columnas de Política/Opinión en El Espectador, El Tiempo y las columnas de la Silla Vacía y la Silla Vacía Académica, lo cual sugiere que en este tipo de columnas y fuentes de información los temas no se centran en conversaciones relacionadas a SIIPO o en algunos casos pueden ser generales como para poder determinar que el contenido de los artículos de noticias, como lo puede ser la palabra "Gobierno". En este orden de ideas, vemos que el contenido de los artículos se fundamenta principalmente en temas relacionados al gobierno, lo cual es esperado al extraer el contenido de los artículos de Opinión y Política, por otra parte, otros de los principales temas más tratados, son sobre la pandemia, educación y sexualidad para el periodo de tiempo que se extrajeron los artículos. De esta manera, como primera hipótesis parece ser que no existe contenido relacionado a SIIPO en los Portales de noticias. Sin embargo, más adelante lo evaluaremos analizando las palabras compuestas más frecuentes y la frecuencia de aparición en el tiempo del término SIIPO y Acuerdos de Paz.



Gráfico 9: Palabras compuestas más frecuentes en Artículos de Noticias



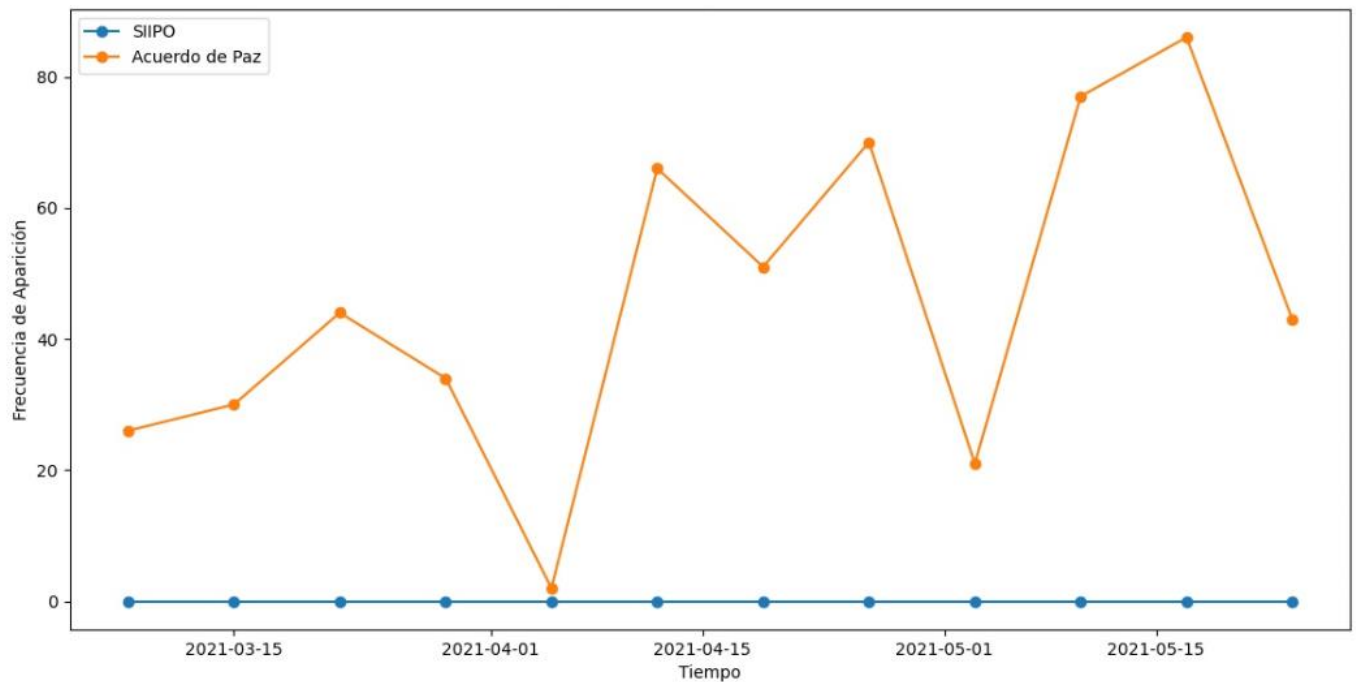
Fuente: Elaboración Propia

En el Gráfico 9, las palabras compuestas de los artículos de noticias se focalizan principalmente en la palabra “Educación Sexual”, pasando a temas coyunturales como la reforma tributaria de 2021 que en sus inicios generó una ola de protestas y reacción social que ha sido persistente hasta estos días y en base a esto se podría explicar la alta frecuencia de esta palabra compuesta.

Por otra parte, la palabra “Derechos Humanos” que puede ser consistente con el sistema SIPO probablemente pueda estar asociada también a las continuas denuncias de derechos humanos por parte de la ciudadanía y la ola de violencia que han generado las protestas del paro nacional que inició a finales de Mayo de 2021, por lo que a nivel global en las palabras compuestas podemos concluir que las columnas de Política/Opinión y académicas se han centrado en la generación de artículos sobre educación sexual y distintos elementos de índole política, social y económica coyunturales como el Paro Nacional. En este orden de ideas, la hipótesis de que no se tratan temas asociados al Sistema Integrado de Información para el posconflicto SIPO, toma mucha mayor fuerza, por lo que a continuación observaremos la frecuencia de veces que aparecen las palabras SIPO y Acuerdo de Paz en los diferentes artículos de noticias.



Gráfico 10: Frecuencia de aparición de las palabras SIIPO y Acuerdo de Paz en los artículos de noticias

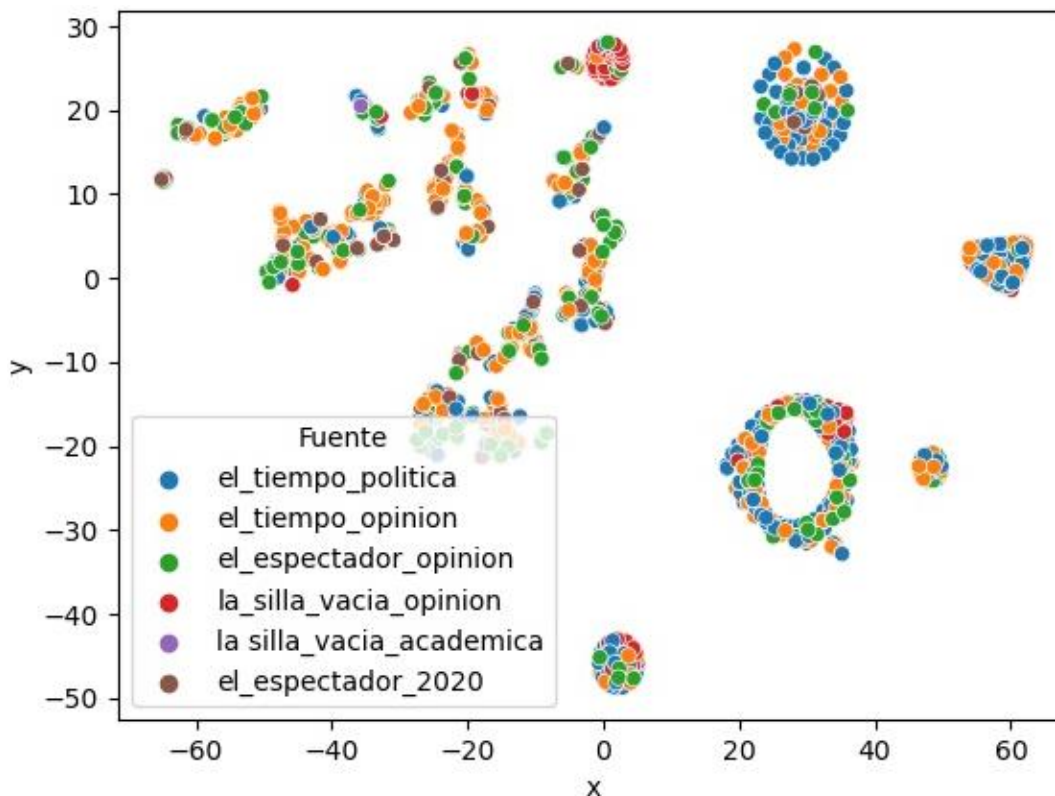


Fuente: Elaboración Propia

Fundamentalmente en el Gráfico 10 podemos observar una tendencia creciente para la palabra “Acuerdo de Paz” que logra alcanzar niveles de aparición en los artículos de noticias hasta por encima de 80 veces. Así mismo, podemos observar que la palabra SIIPO tiene un comportamiento homogéneo a lo largo del tiempo, y que su nivel de frecuencia de aparición es cero para todos los periodos. En este orden de ideas, a pesar de que en las palabras únicas y compuestas más frecuentes para los artículos de noticias no hablen sobre SIIPO en absoluto, en cierta medida los artículos de noticias tratan desde una perspectiva de opinión política y académica sobre los Acuerdos de Paz en Colombia para los meses de marzo a mayo, por lo que el sistema Integrado de Información para el Posconflicto no tiene influenciadores en los portales de noticias utilizados para la extracción de estos artículos.

Otro elemento relevante que nos permite obtener conclusiones sobre el comportamiento y contenido de los artículos de noticias, se presenta a continuación en el Gráfico 11, donde podemos observar el nivel de relación que existe entre las diferentes palabras de estos artículos, es decir si hay relación o no entre estas al igual que su intensidad, de tal manera que nos permita entender de manera gráfica las palabras que se constituyen de forma compuesta, es decir que se escriben juntas de forma frecuente en los artículos de noticias.

Gráfico 12: t-SNE para los artículos de noticias



Fuente: Elaboración Propia

A partir del Gráfico 12 podemos ver que se generan diversos clústeres o grupos, de tal manera que, si se filtran estos grupos por portal de noticia, los resultados son transversales y no existe un patrón reconocible por portal de noticias. Esto quiere decir que a pesar de que son medios distintos, las temáticas utilizadas con respecto a los acuerdos de paz y SIPO son los mismos y no existe gran diferencia entre un medio y otro. Sin embargo, la distancia entre los diferentes clústeres no es demasiado grande, por lo que algunas palabras pueden estar fuertemente relacionadas o ser similares, que explicaría la parte de palabras de alta frecuencia en unigramas que aparecen también como palabras compuestas en las palabras de alta frecuencia para los bigramas en los artículos de noticias.

A partir de esto las palabras agrupadas en los clústeres, se pueden observar a continuación en las nubes de palabras descritas en el Gráfico 13 y 14, de esta manera para entender de forma intuitiva los gráficos, tenemos que entre mayor frecuencia de aparición tiene la palabra mayor será su tamaño, en este orden de ideas permite entender un poco mejor los resultados a nivel general para los artículos de noticias, por tal motivo, palabras como “Gobierno”, “Nacional”, “Personas”, “Pandemia”, “Política” y para bigramas “Educación Sexualidad”, “Paro Nacional”, “Reforma Tributaria” son palabras de alta frecuencia que se ven reflejadas dentro de los clústeres.



5. Conclusiones y recomendaciones

A partir de la metodología desarrollada y de los resultados obtenidos para cumplir los objetivos de este proyecto, planteados en el plan de trabajo, se presentan a continuación las principales conclusiones obtenidas por el equipo de la UCD y las principales recomendaciones para un mejor uso y aprovechamiento del proyecto.

- Las tendencias de contenido de los tweets acerca del Sistema de Información para el Posconflicto SIPO se concentran en opiniones que desean rechazar o destruir los Acuerdos de Paz realizados en Colombia, también parecen incitar hacia la protesta ciudadana. Esto se deduce de palabras de alta frecuencia como “hacer trizas” que cuando se encuentran las combinaciones de 3 palabras más comunes en los tweets la palabra compuesta más frecuente es “hacer trizas acuerdo”, siendo opiniones hacia los acuerdos de paz, mas no al sistema SIPO.
- Las tendencias de contenido de noticias tienen un enfoque más global y no tan centrado al considerar elementos relacionados en otras temáticas como “Educación Sexual”, que es la palabra con mayor frecuencia dentro de los artículos de noticias y fenómenos de coyuntura económica como “Reforma Tributaria” al igual que “Paro Nacional”. Sin embargo, a pesar de que una de las palabras con mayor frecuencia es “Derechos Humanos” en base al resto de palabras comunes que aparecen se considera que esta también se encuentra en el contexto del Paro Nacional, respecto a las altas denuncias de violación de derechos humanos y la reciente intervención de la Comisión Interamericana de Derechos Humanos y su reunión con el gobierno y sectores sociales para evaluar la posible violación de estos por parte del Estado.
- Tenemos que, para tweets y artículos de noticias, no existe una convergencia de discusión hacia el sistema SIPO, esto al considerar la poca mención que tiene este término con frecuencias que se mantienen en valores bajos y cercanos a cero en ambos casos. Para Twitter, la palabra Acuerdo de Paz tiene un mayor número de menciones superando un máximo de 250 menciones, por lo que el contenido de los tweets se centra en los Acuerdos de Paz y no en SIPO. Para el caso de los artículos de noticias existe el mismo patrón de comportamiento, la palabra SIPO no se menciona en absoluto a lo largo del tiempo, mientras que la palabra llega a un máximo de más de 80 menciones en los artículos, a pesar de eso la alta frecuencia de palabras únicas y compuestas no se asocian al Acuerdo de Paz, sino a temas coyunturales en términos de política, economía y sociales en Colombia.
- Los principales influenciadores en Twitter que generan contenido relacionado al sistema SIPO son participantes de la política colombiana y el activismo político, como Rodrigo Londoño el segundo mayor influenciador en Twitter, después de Manuela Marín que también participa en la Dirección Distrital de Bogotá. Por otra parte, también existen cuentas asociadas a grupos del Estado que son grandes Influenciadores como la cuenta “Defendamos la Paz”. Finalmente, a pesar de que la mayor influenciadora en Twitter es Manuela Marín el resto de los influenciadores tienen una cantidad de Retweets similares, por lo que existe una distribución cercana a la uniforme para los Retweets, de tal manera que los principales diez mayores influenciadores de contenido sobre SIPO en Twitter tienen niveles de influencia similares.
- Por otro lado, no se encontraron influenciadores fuertes que permitan generar tendencias sobre SIPO dentro del periodo de tiempo analizado en Twitter.

6. Socialización

El proyecto principalmente ha sido socializado con el Grupo de Seguimiento de Paz y Estabilización; Dirección General DSEPP y Grupo de Proyectos Especiales, Equipo de Paz y Víctimas.



Contacto

Si tiene alguna duda, comentario o sugerencia sobre este proyecto, o si le gustaría conversar con la Unidad de Científicos de Datos sobre la posibilidad de una nueva fase para el mismo, puede comunicarse con nosotros a través del correo electrón