

**Karines RODRÍGUEZ-DÍAZ**

Universidad de Oriente. Cuba. karines@uo.edu.cu

**Dra. Yamile HABER-GUERRA**

Universidad de Oriente. Cuba. yhaber@enet.cu

## **Análisis de sentimientos en Twitter aplicado al #impeachment de Donald Trump**

### ***Sentiment analysis on Twitter applied to Donald Trump's #impeachment***

**Fechas** | Recepción: 02/03/2020 - Revisión: 08/06/2020 - En edición: 24/06/2020 - Publicación final: 01/07/2020

#### **Resumen**

El presente artículo expone conclusiones sobre los sentimientos emitidos en Twitter con respecto al *impeachment* de Donald Trump. Asumimos un enfoque mixto, de tipo descriptivo, que permitió la obtención de resultados sobre aspectos emotivos asociados al discurso de usuarios de Estados Unidos en la esfera pública. La perspectiva investigativa utilizada reside en una estrategia transdisciplinaria para entender los sentimientos relevantes de los tweets en el contexto de su producción. En el caso de estudio se aplican metodologías de minería de datos y sus tipologías: texto, grafo y multimedia, así como el análisis del discurso multimodal para la extracción y evaluación de valores afectivos respecto al juicio. Proponemos tener en cuenta todos los soportes, códigos y prácticas de producción discursiva en el microblogging: los textos escritos e hipertextos, enlaces sociales, reacciones, contenidos multimediales y las identidades de los usuarios en una relación sistémica de significados. Las publicaciones de las comunidades virtuales en las redes sociales sobre procesos políticos son de vital estudio en situaciones electorales y acontecimientos mediáticos de alto grado de polarización.

#### **Palabras clave**

Esfera pública; impeachment; comunidades virtuales; análisis de los sentimientos; análisis del discurso multimodal; minería de datos

#### **Abstract**

*This article presents conclusions around the sentiments expressed on Twitter concerning Donald Trump's impeachment. A mixed approach, of a descriptive type, enabled us to obtain results on emotional aspects associated with users' discourse from the United States in the public sphere. The perspective employed was a transdisciplinary strategy to understand the relevant feelings of tweets in the context of their production. In this case study, data-mining methodologies and their typologies were applied: text, graph and multimedia, as well as an analysis of multimodal discourse for the extraction and evaluation of affective values concerning the trial. We took into consideration all supports, codes and practices of discursive production in microblogging: written and hypertext texts, social links and reactions, multimedia content and users' identities in a systemic relation of meanings. The publications of virtual communities in social networks about political processes are important to study in electoral situations and media events with a high degree of polarisation.*

#### **Keywords**

*Public sphere; impeachment; virtual communities; sentiment analysis; multimodal discourse analysis; data mining*

## 1. Introducción

Analizar las opiniones de la esfera pública e incluso las opiniones publicadas es un proceso complejo que inevitablemente se estructura como una visión parcial, aunque útil, en contextos comunicativos. Las redes sociales, como avance tecnológico, no sustituyen a la plaza pública, magnifican la percepción de democracia, aunque al menos dan paso a la participación ciudadana de usuarios en las conversaciones sobre temas públicos. Las audiencias de los medios de comunicación generan sentido y emociones, usualmente, sobre contenidos que han tratado los medios o instituciones con mayor opulencia comunicativa. Los usuarios distribuyen argumentos que hacen alusión emotiva, ya sea de manera directa (citando la fuente) o indirectamente en sus publicaciones, a contenidos de las noticias oficiales, así lo explica los argumentos teóricos sobre el *gawatching* (Burns, 2011; Burns y Highfield, 2015).

Se vive en una sociedad en Red donde los públicos producen y reproducen sus capacidades de creación (Martínez, 2013), en este contexto es más fácil estudiar sus opiniones respecto a sucesos de polarización política porque accedemos a lo que constituye una opinión publicada, un discurso público. Podemos encontrar a los electores en las plataformas sociales participando, comentando de manera instantánea y reticular con contenidos a los que pueden reaccionar los demás usuarios aprobando, distribuyendo, negando, y/o refutando en las comunidades virtuales<sup>[1]</sup>. Aunque según Enguix (2017), "la gran aportación de las redes sociales a los medios de comunicación digitales o analógicos es el aumento innegable de la difusión de sus publicaciones."

Numerosos autores han valorado el potencial de Twitter como espacio para la comunicación política (Moya y Herrera, 2015; Pérez, Haber, Díaz y Zamora, 2017; Gómez, 2014). Según Orihuela (2011) Twitter y las redes sociales están llevando a la gente corriente la posibilidad de tener voz pública. Engesser y Humprecht (2015) destacan a Twitter como una plataforma en la cual periodistas y políticos son especialmente activos. El servicio de microblogging ocupa el número 13 de preferencia a nivel mundial cuenta con alrededor de 340 millones de usuarios activos<sup>[2]</sup> y es uno de los espacios digitales más usados con fines políticos y propagandísticos, así lo demostró la campaña de Donald Trump que le valió el triunfo de su primer mandato. De igual forma Jack Dorsey, fundador de Twitter, al anunciar las restricciones a los anuncios políticos en esta red social manifestó el papel que la misma ha jugado en contextos electorales y su influencia sobre los grupos virtuales<sup>[3]</sup>.

Ahora bien, en el análisis de la voz pública sobre procesos que incluyen la construcción de una identidad colectiva basada en emociones, procesos culturales, identitarios, simbólicos, presiones sociales y fuentes a las que han sido expuestas dichas categorías trascienden, en la mayoría de los casos, al contenido factual de posible estudio. Así, la opinión pública está compuesta por diferentes enfoques válidos: los intereses y perspectivas de los emisores, los interventores y los receptores (Briones, 2018); la articulación de su discurso se encuentra atravesada por tres dimensiones: lo cognitivo, lo social y lo sistémico (Portillo, 2004). Pero, de manera funcional, es posible analizar el estado de opinión pública en el microblogging a través de objetos auditables: el discurso, sus características, prácticas y las relaciones sociales que de este se derivan.

La producción científica actual asociada a la extracción de emociones publicadas por usuarios en Twitter se centra en el análisis de los sentimientos como metodología efectiva basada en minería de textos sobre big data<sup>[4]</sup>. Aunque a nuestra consideración debiese tomar en cuenta la evaluación del significado emotivo en los contenidos multisoportes y los entramados sociales que representan simbólicamente cuestiones afectivas en el discurso.

En tal sentido el presente artículo expone la utilización de un enfoque mixto basado en el uso de varios tipos de minería de datos y del análisis del discurso multimodal como metodología. La elección de dicho entramado metodológico sugiere evaluar de manera generalizada el contenido emotivo sin segmentar la muestra a un soporte específico e incluyendo categorías como las relaciones sociales, influencia social, los enlaces y el análisis de los sentimientos en archivos multimedia. Presentamos los resultados de la aplicación de estas técnicas y métodos de investigación en un caso de estudio sobre la producción discursiva que con el *hashtag* *#impeachment* publicaron usuarios de Estados Unidos en la plataforma de Twitter.

### 1.1 Precedentes investigativos

"El análisis de sentimientos (*sentiment analysis*, SA) es una de las principales técnicas de estudio de datos textuales a gran escala (*big data*) empleadas en la investigación en ciencias sociales y en comunicación política. Su objetivo es reconocer y evaluar el valor emocional existente detrás de los textos analizados, a

través de su estructura, clasificándolos en positivos, negativos o neutros. En la actualidad, esta metodología se aplica principalmente en la interpretación de los textos difundidos en medios sociales como Twitter” (Arcila, Ortega, Jiménez y Trullenque, 2017: 975) y con un enfoque asociado a atributos léxico sintácticos (Vilariño et al, 2015) que permiten cualificar el texto.

Al análisis de sentimientos se le conoce también con el nombre de minería de opiniones. Otros autores lo denominan análisis de subjetividad, la materia posee conexiones con la informática afectiva, el reconocimiento computacional y la expresión de la emoción. Como campo de investigación, es una parte de la lingüística computacional, área que estudia el procesamiento del lenguaje natural y la minería de textos cuyo objeto es el discurso (Fornari, Abeille, Ferrero, Pérez, y Boglione, 2019).

“A diferencia del análisis de sentimientos manual con codificadores humanos o del análisis automático asistido por ordenador el uso de la minería de textos utiliza procedimientos del aprendizaje automático supervisado para generar modelos basados en datos previamente etiquetados y así poder predecir con un significativo grado de fiabilidad el sentimiento de los mensajes, además permite ejecutar filtros sobre las consultas, por ejemplo, de fechas, idiomas, lugares geográficos o etiquetas incluidas en el texto a analizar del mensaje” (Arcila et al, 2017).

Las líneas de investigación en espacios de comunidades virtuales, donde confluyen gran cantidad de datos de acceso abierto para obtener resultados sobre emociones y perspectivas públicas, han girado en cuatro aspectos esenciales:

- Propuestas teórico metodológicas para el uso de softwares, algoritmos<sup>[5]</sup> y enfoques propios de ciencias cibernético-informáticas para la detección de patrones enfocados en sentimientos de los usuarios (Colle 2002, 2013, 2017; Daniel, 2010; Cohen y Hamilton 2011; Mancera y Pano, 2013a, 2013b; Diakopoulous, 2015; Coddington, 2015; Pérez, Haber y Duvergel, 2016; Pérez et al, 2017; Hermida y Young, 2017; Vállés y Codina, 2018; Conde, Pullaguari y Prada, 2019).
- Uso de la minería de textos (Zappavingia, 2011; Reyes, 2012; Ortega, Fonseca, Gutiérrez y Montoyo, 2013; Vilares, Alonso y Gómez, 2013; Abascal, López, y Zepeda, 2014; Almgrem y Olsson, 2015; Vilariño et al, 2015; Arcila et al, 2017, Figueira, y Guimarães, 2017; Verbeke, Berendt, d'Haenens, y Opgenhaffen, 2017; Hernández, 2017; Van Hee, 2017; Kwabla, Kwame, y Katsriku, 2017; Martínez, 2017; Reyes, Paniagua y Sánchez, 2017; González, Hurtado y Pla, 2018; García, Henríquez y Herrera, 2019).
- Análisis cualitativo para la detección de valores afectivo emocionales (Gómez, 2014; Ventura, 2016; Marín y Quintero, 2018).
- Propuesta de uso de herramientas o técnicas mixtas (cuantitativas y cualitativas) para la detección de sentimientos y afectos (De Uribe, Pascual, y Gascón, 2016; Suau, Percastre, Palá y Pont, 2017; Baviera, 2017; Rodríguez y Haber, 2017; Aguirre, Hernández, Briceño y Marín, 2018; Vallejo, 2018).

En las propuestas teórico metodológicas vale destacar la producción científica de Daniel (2010) quien expone metodologías de varias ciencias incluyendo perspectivas y paradigmas de análisis cualitativo, cuantitativo y de enfoque mixto lo que supone un análisis transdisciplinar para el estudio de comunidades virtuales. También Colle (2017) expone la necesidad de algoritmos que, aunque no son 100 % efectivos, permiten detectar patrones y cualificar el contexto en grandes corpus de datos. El uso de la minería de textos para la detección de polaridad y opiniones es bastante efectivo, aunque está en constante mejoramiento en base a sus características lingüísticas y aún precisa de algoritmos más efectivos en la detección de recursos como la ironía y el sarcasmo aseveran autores como: Reyes et al (2012), Ortega et al. (2013), Hernández (2017) y Van Hee (2017).

En la minería de textos se utilizan técnicas simbólicas que hacen uso de recursos léxicos para revisar la polaridad. Kwabla et al. (2017) explica que como procesamiento se emplean la lingüística de corpus y la minería basada en concordancia de texto o clustering<sup>[6]</sup>. La tarea de identificar el sentimiento predominante en un texto escrito es una labor compleja incluso para un ser humano por este motivo, el análisis de sentimientos automatizado requiere un desarrollo y perfeccionamiento deliberado continuo que ha sido planteado desde dos enfoques: aproximaciones semánticas y técnicas de aprendizaje computacional.

Sin embargo en la minería de textos se responde al filtrado de sentimientos en un soporte específico dentro del constructo simbólico de la muestra analizada: el texto escrito, sus características y cualidades. El texto escrito sigue siendo la parte importante de la comunicación de la cual no podemos prescindir, su extracción y calificación da paso a que softwares especializados sean capaces de medir la polaridad<sup>[7]</sup>, los enlaces y las reacciones en grandes corpus de datos. Sin embargo, ante las propiedades de las publicaciones del usuario en las redes sociales, no podemos eliminar en el contexto enfocado hacia la

detección de afectos categorías como las relaciones sociales y el contenido multimedia en la expresión de emociones.

El tweet es una unidad de sentido de 280 caracteres que no puede ser separada de sus retweets, me gusta, videos, imágenes, o enlaces que forman parte del contenido, de igual forma han de tenerse en cuenta el número de respuestas como patrones de reacción, aunque estas pueden ser estudiadas separadas del texto rey (o tweet principal) pues proceden de otro usuario. Las reacciones (me gusta y retweet) y respuestas dentro del tweet aunque son atributos medibles a través del filtro de textos son el resultado directo de la articulación de relaciones sociales. Según Cansino et al (2016: 16).

“(…) el tweet es persuasión. Un tweet que nadie lee es un tweet que nunca existió. Para trascender, un tweet tiene que seducir, persuadir, convencer, motivar... Sólo así destacará de entre millones de tweets que nacen y mueren cada segundo (...)”

Aunque la mayoría de las investigaciones precedentes se orientan hacia la minería de textos dentro de la minería de opiniones existen pocas excepciones que le han prestado interés a las relaciones sociales y el contenido multimedia. En tal sentido autores como Welbers y Opgenhaffen (2019), Wu, Hofman, Mason y Watts (2011), y Gruz, Wellman, y Takhteyev (2011) denotan una particular importancia a los enlaces sociales en la conversación en comunidades virtuales como factor determinante en la producción de sentido y los discursos emotivos. Los enlaces sociales son medibles a través de la minería de grafos<sup>[8]</sup>, técnica que proporciona herramientas para mapear las estructuras de datos y para encontrar nuevas conexiones (relaciones) entre los objetos o nodos (Chakrabarti, Papadimitriou, Modha, y Faloutsos, 2004).

En cuanto al análisis del contenido multimedia Conde et al (2019) realizan una comparativa cuantitativa de la producción en Twitter que menciona la multimedialidad de mensajes políticos publicados por candidatos presidenciales en España, Ecuador y Colombia, prestando atención a los mensajes en diferentes soportes en el contexto del microblogging, aunque los autores no profundizan en el contenido emotivo del mismo. Además, Vásquez (2019) habla de un análisis multimodal en la relación semiótica entre texto del tweet y la imagen añadida, suscribimos que dicha relación supone una unidad de sentido.

Lo que llamamos emociones, sentimientos y opiniones desde el punto de vista de la interpretación no solo puede ser determinado a través del análisis del texto escrito en la minería de textos pues supone una visión segmentada de los sentimientos expresados en la esfera pública. Ha de ser contemplado la minería de grafos y multimedia, las identidades y los factores de influencia social para su determinación contextual. La minería multimedia es útil para la categorización basada en la forma como: tipo de archivo, duración, tamaño, etcétera. No obstante, es menos efectiva en el análisis del contenido, por ello es permisible la aplicación de metodologías cualitativas de tipo manual para determinar los afectos utilizados en soportes de audio, video y contenido interactivo.

Estudiar no solo el texto escrito sino incluir valores multimodales es importante dado que “no es en los datos estructurados en donde se encuentra el mayor volumen de información que existe al nivel mundial, sino en multimedia como es el caso de imágenes, texto, audio y vídeos” (Oviedo y Velez, 2016:127).

La minería de imágenes, comúnmente, realiza una caracterización por medio de histogramas de color, se pueden extraer características de textura y forma. La minería de audio efectúa un aventajamiento para extraer características de bajo nivel en el dominio del tiempo y la frecuencia. Finalmente, la minería de vídeos se presenta como la conjunción de características de textos, imágenes y audio. Se puede identificar que la minería multimedia presenta problemas de dimensionamiento, ya que es representada por vectores de características en alta dimensión, lo que afecta el desempeño de métodos de analítica basados en el contenido, por tanto, es pertinente un análisis cualitativo para la extracción de datos emotivos en imágenes, audio y video. Arnáiz y Filardo (2020) expone un enfoque multimodal en el estudio de imágenes en Twitter.

Lo que proponemos no es una superación de los argumentos y tesis de los autores citados anteriormente, abogamos por una inclusión de sus exposiciones en una perspectiva más abarcadora. Dada su naturaleza, la mayoría de las investigaciones expuestas en este apartado y probablemente por el objetivo que perseguían los investigadores, han segmentado el análisis de sentimientos y opiniones a un soporte específico. Los estudios asociados al discurso, espacio productivo donde encontramos la representación de los sentimientos, han de utilizar metodologías múltiples para la detección de patrones en sus muestras multisoportes. Así mismo precisan de softwares completos cuyas facilidades detecten elementos afectivos en textos, grafos y archivos mutimedias. El investigador ha de presentar una actitud activa en la revisión de los filtros y la detección de la efectividad de los algoritmos, un acompañamiento humano.

De igual forma recomendamos no suprimir la utilización de técnicas cualitativas como el análisis del discurso tan necesaria en las investigaciones de los materiales audiovisuales cuya expresión y contenido no puede ser categorizada eficazmente aún por un software de *data mining*. Hablamos de una asimilación de las metodologías expuestas por los autores, todas útiles, perspectivas cualitativas y cuantitativas, pero más eficientes juntas en un enfoque mixto, aunque complejo y transdisciplinar. Ahora bien, planteamos una visión generalizada que integre varias categorías o códigos propios de las prácticas discursivas de Twitter: textos (tweets, respuestas, enlaces, hashtags, identidades de usuarios disponibles en la bio), grafos (relaciones sociales y de etiquetas), multimedia (contenido y forma en formatos de fotos, audio y video) y las relaciones texto-grafo-multimedia. Un sistema simbólico ha de ser analizado a partir de las partes que componen el conjunto de sus prácticas de significado y producción emotiva.

Por tanto, en el presente reporte exponemos una investigación descriptiva sobre las emociones publicadas por los usuarios en comunidades virtuales de Twitter. Desde un enfoque mixto que utiliza herramientas cuantitativas y cualitativas para la obtención de resultados, nos enfocamos en el estudio de caso particular del *#impeachment* en el microblogging.

## 1.2 Contexto político del *#impeachment*

El desarrollo del proceso de *impeachment* de Donald J. Trump, presidente de los Estados Unidos, tuvo inicio formal el 18 de diciembre de 2019 cuando tras las escuchas testimoniales, la Cámara de Representantes aprobó el acusar al presidente de obstrucción al Congreso y abuso de poder. La primera sesión se celebró el martes 21 de enero y la última el 5 de febrero de 2020 cuando el senado absolvió y cerró el proceso luego de no utilizar testigos, pruebas ni documentación.

La noticia tuvo un gran impacto en la palestra internacional dada la cercanía de las elecciones de noviembre de 2020 en Estados Unidos y se encontró entre los acontecimientos más seguidos por los usuarios que produjeron sentido durante el trascurso. El presente reporte de investigación ahondó en las emociones a partir de su representación simbólica <sup>[9]</sup> en el discurso. Teniendo en consideración las prácticas sociales realizadas durante el juicio a Donald Trump en Twitter utilizamos la minería de datos, sus diversas tipologías, y el análisis del discurso multimodal sobre los tweets sustraídos bajo el hashtag *#impeachment* y cuya geolocalización se ubicó en Estados Unidos. Las publicaciones analizadas fueron publicados entre el 21 y el 25 de enero de 2020, se seleccionó dicha fecha a fin de valorar los contenidos emotivos a inicios del juicio político, el período fue además representativo dado la prevalencia de tweets originales con respecto a respuestas, menciones y retweets respectivamente.

Se tuvieron en cuenta los tweets publicados que no procedían de instituciones mediáticas a fin de evaluar el estado de emociones de la esfera pública activa que generó contenidos en Twitter durante la fecha. En la investigación la minería de datos y el análisis del discurso multimodal se comportan como herramientas útiles para un análisis de sentimientos y opiniones en el contexto del microblogging. De igual forma se describe la significación que para la esfera pública tuvo el *#impeachment* desde la perspectiva de la comunidad virtual que produjo significados al respecto siguiendo la etiqueta. Por tanto, en el estudio, se responden las siguientes preguntas de investigación:

- ¿Qué polaridad asumieron los usuarios implicados en el discurso sobre el *impeachment*?
- ¿Cuáles fueron las emociones con mayor número de reacciones asociadas al *impeachment*?
- ¿Cuáles fueron los puntos de encuadre, consideraciones o estados de opinión con respecto al inicio del juicio de *impeachment* de Donald Trump?

Las anteriores preguntas de investigación determinaron los siguientes objetivos:

- Comparar la polaridad del total de tweets publicados a fin de medir el estado de opinión de los usuarios sobre el *impeachment*.
- Determinar las emociones en el discurso de los tweets con mayor número de reacciones (retweets, me gusta, respuesta).
- Identificar el encuadre, las consideraciones o estados de opinión que dieron los usuarios de los tweets al inicio del *impeachment* de Donald Trump.

## 2. Metodología

El presente artículo propone la utilización de la minería de datos y sus diversas tipologías (minería de textos, minería de grafos y minería multimedia) para el análisis de sentimientos en el contexto de Twitter, así mismo incorpora la metodología emergente de Análisis del discurso multimodal como herramienta cualitativa en un contexto discursivo multimedia. La propuesta metodológica acá presente, estructurada en un estudio de caso, asume una postura contextualizada en la producción de sentido emotivo en sus diversos códigos en el *microblogging*: contenido multimedia, identidades de los usuarios, relaciones de influencia y relaciones entre usuarios, relaciones intertextuales, de distribución, de etiquetas, reacciones, entre otras. Debemos ver el discurso producido como una interfaz que representa un contenido simbólico, según Scolari (2018) en la 9na ley de las interfaces: la interfaz es un ecosistema que incluye relaciones de exclusión, inclusión, convergencia, divergencia, sustitución y extensión –reducción, que a consideración de las autoras pueden identificarse en la detección de patrones por parte de los algoritmos que incluye elementos categoriales susceptibles de estudiar en los códigos mencionados. Por lo tanto, promovemos la inclusión de las relaciones sociales y el análisis de archivos multimedia para detectar los sentimientos sin dejar ninguna de las categorías antes expuestas fuera del asunto.

Para la obtención de resultados la minería de datos se aplicó con el software Stela, desarrollado por la empresa Datys en Cuba. La extracción de los tweets se realizó desde el 21 de enero y hasta 7 días después del 25 de enero a fin de descargar las reacciones que tuvieron los tweets publicados con el hashtag *#impeachment* durante el período analizado. Para la extracción de los tweets se utilizó la interfaz de programa de aplicación (API) de Twitter (REST y Streaming), recurso libremente disponible.

El API REST permite descargar y filtrar el histórico de mensajes de los últimos 7 días de manera gratuita.

El análisis de la polaridad con el algoritmo SSA-UO del software Stela presenta un 0,06 % de margen de error en detecciones de polaridad lo cual es bastante aceptable para la identificación de cualidades en grandes corpus de datos, tarea casi imposible de manera manual. El software Stela procesa los mensajes a través de varios tipos de minería: textos, estructura (enlaces), grafos y multimedia, lo que permite una clasificación generalizada. Así mismo posee filtros para detectar y cualificar los tweets según caracteres y propiedades específicas.

Según Baeza (2009) la minería de datos en redes sociales funciona con utilidad en la determinación de percepciones. El concepto de minería de datos o data mining agrupa técnicas computacionales que permiten descubrir informaciones, especialmente características que relacionan de un modo inesperado – o difícil de descubrir – los valores de múltiples variables en una gran cantidad de registros. Los métodos de *data mining* revelan esta información y la transforman en un valioso conocimiento tanto retrospectivo (histórico) como prospectivo (proyecciones) o "comprensivo" (entender lo que ocurre), siendo así muy importante para las tomas de decisiones en las empresas, organizaciones y gobiernos. Por lo tanto, la minería de datos es por esencia una metodología de exploración y descubrimiento.

La minería de datos no elimina la participación humana para resolver la tarea completamente, pero simplifica significativamente el trabajo y permite a un analista que no es un profesional de las estadísticas y de la programación administrar el proceso de extraer conocimiento a partir de los datos. Los métodos son variados y el software existente incluye habitualmente una "batería" de programas que operan de distintas maneras y entregan distintos tipos de resultados, en su mayoría acompañados de formas visuales destinadas a poner en evidencia las relaciones descubiertas. Las operaciones que es posible realizar son, por ejemplo, la clasificación, la estimación, la predicción, la detección de relaciones, el modelamiento, la segmentación (clustering) y la detección de desviaciones (Colle, 2017).

Los algoritmos automatizados desempeñan un papel muy grande en algunos servicios. En este sentido, el objetivo de la minería de datos es aprovechar la hiperabundancia de información. Por eso para el trabajo en esta técnica es muy relevante establecer términos de búsqueda concretos. El programa Stela de minería de datos en Twitter es particularmente útil por su amplia gama de posibilidades para extracción y filtrado, aunque es menos funcional factualmente en la descripción del contenido de imágenes, audio, videos y soportes interactivos como las encuestas. Por tanto, se precisó el análisis de discurso multimodal para la determinación de la relación entre el texto escrito y el contenido en otros soportes.

"El análisis del discurso multimodal constituye un paradigma emergente en el campo de los estudios del discurso que amplía el estudio del lenguaje *per se* al estudio del lenguaje en combinación con otros recursos" (O'Halloran, 2012: 75). Ello implica que el análisis del discurso multimodal sea mucho más abierto a la variedad de soportes y a la multimedialidad del discurso en plataformas digitales como Twitter, dado que incluye varios recursos semióticos como la imagen, los videos y los valores reacción añadidos al texto en el microblogging como el retweet, me gusta, y respuestas. También se tuvo en cuenta el hecho de que

los usuarios utilizaran argumentaciones para validar el enfoque emotivo con respecto al inicio del *impeachment* en el corpus analizado, aunque no lo hicieran desde un modo clásico de argumentación como comprueba Ventura (2016) que se produce en Twitter.

El análisis del discurso multimodal nos permite visualizar el significado de un acontecimiento en específico. La teoría de las metafunciones modela el potencial de significado de los recursos semióticos en: - el significado ideativo (es decir nuestra concepción del mundo) involucra: lo experiencial, la representación y la forma en que retratamos la experiencia obtenida, - el lógico; - el interpersonal: la puesta en acción de las relaciones sociales y el -textual: la organización en forma de textos y unidades coherentes (ledema, 2003). Dichas metafunciones en el microblogging son susceptibles de medir a partir de los datos contextuales sustraídos.

Estudiaremos el constructo del significado textual entendiendo como textos los tweets emitidos, teniendo en cuenta que no solo suponen una dimensión emotiva personal, sino que son el resultado de las fuentes a las que han sido expuestas los usuarios y las presiones de grupo producto de las relaciones sociales. A fin de cuentas, según Foucault (1973) los discursos instituyen, ordenan y organizan la interpretación sobre la sociedad, las prácticas sociales, los actores sociales y las relaciones entre ellos, mediante la construcción de versiones (representaciones simbólicas), que contienen valores, opiniones, etcétera. En tal sentido descubrimos órdenes, reglas y regularidades en el discurso, y se reconoce los encuadres adoptados por la esfera pública.

Es por ello que en el contexto de la presente investigación nos basamos en un análisis global y luego en una guía de análisis fino (Wodak, 2003) sobre la comunidad virtual, entendida como el número total de usuarios que de manera activa publicaron contenido o reaccionaron al contenido publicado dando retweet, me gusta o respondiendo con la etiqueta *#impeachment* en Estados Unidos.

La guía de análisis global se aplicó sobre el corpus total de 13 661 tweets para determinar las formas del discurso más utilizadas en el contexto del microblogging: tweets, retweets, respuestas y menciones, la medición comparativa de las polaridades a las que respondían, las referencias a contenidos externos a través de hipertexto, las interrelaciones entre el hashtag principal y otras categorías textuales repetidas en el discurso, además del uso de recursos como el video y la imagen para la producción de significados.

Por otro lado, el análisis fino se aplicó a una muestra de 100 tweets originales considerados los de mayor influencia social<sup>[10]</sup>. Particularmente se tuvo en cuenta el encuadre de los tweets en el análisis fino porque determinaron un aumento en las reacciones de los usuarios expuestos o emisores y una sincronía en el aspecto emotivo. Para este análisis fino se aplicó la siguiente guía de análisis:

1. Polaridad asumida en la mayor parte de los tweets.
2. Comparación entre el número de me gusta y el número de retweets.
3. Encuadres respecto al *impeachment* que fueron valorados por los usuarios con mayor influencia social (o con mayor número de relaciones sociales en el contexto).
4. Uso de imágenes o videos. Valoración del significado del discurso visual con respecto al texto escrito en el *tweet*.
5. En el punto número 2 diferenciamos las reacciones entre número de me gusta y de *retweets* dado que desde el punto de vista del significado dar me gusta implica un acercamiento emotivo o aceptación primaria, pero el *retweet* funciona de manera más directa- como una doble aceptación o distribución- pues implica una interrelación mayor, cita textualmente un mensaje, lo dota de mayor fuerza y apoyo dentro del discurso de la esfera pública.

### 3. Análisis de los resultados

Como parte del análisis global se determinaron las formas del discurso utilizadas en el corpus de la muestra (como ilustra la Tabla 1). La herramienta de minería de datos Stela, utilizada para determinarlo, facilita estos resultados con el filtro análisis de contenido, que muestra un resumen sobre el total. En el caso de estudio es representativo el uso de tweets originales con respecto a las formas de retweets que contiene el menor porcentaje de emitidos. El caso se comporta de manera anómala puesto que en otras investigaciones la conversación se da utilizando mayoritariamente el retweet por lo que tal conclusión denota un grado de participación y publicación de contenido relevante por parte de usuarios.

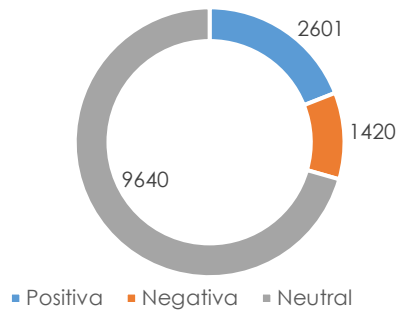
**Tabla 1: Formas del discurso utilizadas en el corpus de tweets de la muestra**

Forma del discurso	Cantidad publicada	% sobre el total emitido
Tweets	4485	33 %
Respuestas	4267	31 %
Menciones	3292	24 %
Retweets	1617	12 %

Fuente: Elaboración propia a partir de resultados obtenidos a través del software Stela.

Ahora bien, las polaridades de las distintas formas del discurso se comportaron mayoritariamente neutrales seguidas de los enunciados positivos siendo el menor número del total los tweets negativos. (como se muestra en la Ilustración 1).

**Ilustración 1: Medición comparativa de las polaridades del total de tweets**



Fuente: Elaboración propia a partir de resultados obtenidos a través del software Stela.

Las referencias a contenidos externos a través de hipertexto, son en su mayoría enlaces a publicaciones de medios de comunicación que argumentan la opinión enunciada en el tweet. Como se expone en la Tabla no. 2 los hipertextos distribuidos en su mayor parte llevan a trabajos periodísticos de CNN, Rolling Stones, Fox News o a eventos de Twitter, aunque otros usuarios también distribuyeron contenidos de estos medios de manera indirecta sin compartir el enlace, tal es el caso del video donde el presidente Donald Trump expresa que tenía todo el material (refiriéndose a las evidencias) mientras que “ellos” no lo tenían (véase <https://bit.ly/3ceLjUd>)

**Tabla 2: Hipertextos (urls) más compartidas en los tweets**

Url compartida en el tweet	Número de tweets que contienen la url
<a href="https://fxn.ws/369XIUX">https://fxn.ws/369XIUX</a>	23
<a href="https://twitter.com/li/events/1220068532370599936">https://twitter.com/li/events/1220068532370599936</a>	21
<a href="https://www.cnn.com/2020/01/22/politics/trump-lawyers-impeachment-false-claims-scif/index.html">https://www.cnn.com/2020/01/22/politics/trump-lawyers-impeachment-false-claims-scif/index.html</a>	18
<a href="https://www.cnn.com/2020/01/23/politics/donald-trump-impeachment-trial-witness/index.html">https://www.cnn.com/2020/01/23/politics/donald-trump-impeachment-trial-witness/index.html</a>	18
<a href="https://twitter.com/li/events/1219696006310199296">https://twitter.com/li/events/1219696006310199296</a>	16
<a href="https://www.rollingstone.com/politics/politics-news/trump-impeachment-evidence-we-have-all-the-material-they-dont-941140/">https://www.rollingstone.com/politics/politics-news/trump-impeachment-evidence-we-have-all-the-material-they-dont-941140/</a>	15
<a href="https://www.oann.com/sen-cruz-impeachment-trial-is-excuse-for-democrats-to-continue-partisan-attacks/">https://www.oann.com/sen-cruz-impeachment-trial-is-excuse-for-democrats-to-continue-partisan-attacks/</a>	15
<a href="https://twitter.com/repvaldemings/status/1220017702011535364">https://twitter.com/repvaldemings/status/1220017702011535364</a>	13
<a href="https://www.foxnews.com/opinion/trumps-senate-impeachment-trial-judge-andrew-napolitano">https://www.foxnews.com/opinion/trumps-senate-impeachment-trial-judge-andrew-napolitano</a>	13

Fuente: Elaboración propia a partir de resultados obtenidos a través del software Stela



Las interrelaciones entre el hashtag principal (*#impeachment*) y otras categorías textuales repetidas en el discurso incluyen el uso de elementos que fueron identificados a través de filtros de concordancia de textos en la interrelación con otras etiquetas y palabras tales como juicio, senado, presidente, y demócratas. Así mismo la conversación se dirigió en reiteradas ocasiones a la figura de Schiff (Miembro-demócrata- de la Cámara de Representantes de Estados Unidos y nombrado Jefe de los fiscales en el juicio de *impeachment*), Pelosi (Presidenta de la Cámara de Representantes de los Estados Unidos. Miembro del Partido Demócrata y considerada uno de los rostros principales en la firma y el proceso de juicio político a Trump), Alan Dershowitz y Ken Starr, ambos abogados de Donald Trump durante el *impeachment*. Además, se mencionó directamente al usuario del presidente Donald Trump (@realdonaldtrump) unas 929 veces durante los 4 días tomados como referencia para este estudio.

El uso de recursos como el video y la imagen se identificó en el 42 % de los *tweets*, en la mayoría de los casos completa el sentido del texto escrito en los 280 caracteres del *tweet* o es anunciado por dicho texto. Como es el caso del siguiente con 18 684 *retweets* (traducido al español por los autores):

@tommalone1961: Nuevo video sobre el *impeachment*. El Senado existe para momentos como este #HoldtheLine

El *tweet* (arriba) presenta y completa su sentido con el video que publica donde se expresa la importancia del senado en casos de confrontación política como el juicio del *impeachment*. De una manera promocional se expresa en 2:17 min el papel preponderante del Senado en la toma de decisiones durante el juicio, acá se muestra cómo @tommalone1961 distribuye el contenido publicado por @NRSC (usuario de *The Senate Majority*), y lo dota de una doble aceptación pues lo suscribe a través del *retweet* de una cuenta verificada de la mayoría conservadora del Senado estadounidense.

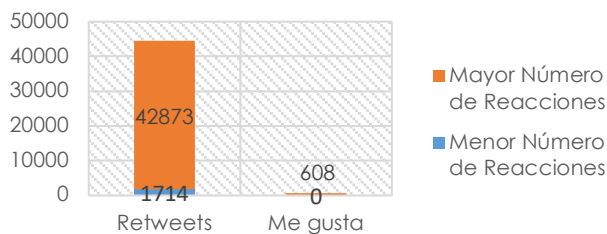
El análisis fino aplicado a 100 *tweets* más influyentes, y respondiendo a la guía de análisis expuesta en la metodología: 1) la mayoría de estos responde a la polaridad neutral, solo en algunos casos como el *tweet* de @tommychong, cuenta verificada (comediante, actor, director y activista en USA), se muestra una actitud negativa, aunque el contenido del *tweet* no es de los más influyentes, solo con 81 *retweets* y 608 me gusta. Expone:

@toomychong: Trump será eliminado ya sea por juicio político, juicio en el Senado o por voto popular. No está mentalmente apto para ser Presidente.

Recibió alrededor de 20 respuestas la mayoría refutando el argumento o bien en contra o aludiendo que en el contexto de las elecciones ganadas por Donald Trump el voto popular no contó.

A continuación, expresamos en el gráfico cómo se comportó las reacciones de la esfera pública entre *retweets* y me gusta, no incluimos las respuestas puesto que el número de ellas en los *tweets* más influyentes no fue representativo. 2) En este caso concluimos que el consenso en los 100 *tweets* evaluados en el análisis fino en la esfera pública de la comunidad virtual que siguió el hashtag *impeachment* se dio fundamentalmente vía *retweet*, es decir a través de cita directa (como podemos comprar en la Ilustración 2).

**Ilustración 2. Mayor y menor número de reacciones por tweet**



Fuente: Elaboración propia a partir de resultados obtenidos a través del software Stela

3) De manera general los *tweets* con mayor influencia social publicados por los usuarios de la comunidad virtual asumieron los siguientes encuadres como juicios de opinión respecto al *impeachment*:

Defienden que el juicio político de Donald Trump no tendrá consecución ni fin alguno y para ello encuadran sus argumentos y utilizan frases como las que se expresan a continuación:

- "Fallo miserable", intento demócrata que no llevará a nada.
- No aprobación de testigos ni pruebas.
- Mayoría republicana del Senado
- El *impeachment* es un gasto de dinero y tiempo para el pueblo americano
- Los demócratas no tienen caso, escucharlos es cosa de gracia.
- El *impeachment* va sobre como Trump ganará las elecciones en 2020
- En el juicio se gasta todo el día sin fin alguno, algunos *retweets* comparten el hecho de que el senador Jim Rish fue el primero en quedar dormido durante el *impeachment*.
- Se habla de una impotencia demócrata utilizando palabras contra Trump de manera innecesaria, se asume una impotencia de Schiff durante el ejercicio.

Otras percepciones compartidas mayormente en el asunto valoran la actitud de Donald Trump como autoritaria y segura dada la presunta tranquilidad del presidente y las palabras distribuidas por los medios donde dice que "honestamente tenemos todo el material. Ellos no tienen el material", en tal sentido vale destacar la influencia del *tweet* de @RepValDemings (Miembro de la Cámara de Representantes de los Estados Unidos. Representante en el Décimo Distrito del Congreso de la Florida) que expresa el modo en que el presidente norteamericano encubrió testigos y documentos al pueblo americano y lo refuerza con el uso del video.

Otros elementos, aunque en menor grado de influencia valorados como patrones de repetición por concordancia de texto y *retweet* fueron:

- El juicio político de Donald Trump muestra a políticos divididos y una contradicción representativa entre ambos partidos. Relevante en el discurso en tal caso es el *tweet* con 2066 *retweets* que expresa: @BrandonStraka: Demócrata de 91 años: "¡Deshágase de esta acusación de *impeachment*! ¡Como demócrata estoy harto de esto! #Walkaway

El *tweet* contiene el video (*retweet* de @Aikens\_Josh) que refuerza la idea en la que el político demócrata de 91 años expresa su cansancio ante el proceso y la división entre partidos.

- Hubo discusiones sobre el argumento de la CNN de que el 51% de los americanos aprobaba condenar a Trump. Varios usuarios compartieron, alrededor de 398 *retweets* el siguiente argumento y llevaron a cabo una encuesta en Twitter haciendo más participativa la publicación. Dicha encuesta dio como resultado con 98 % de aprobación la opción absolver a Trump; aunque se obtuvo a partir de 1096 votos.

El texto del *tweet*, que hace un llamado a la acción, mencionaba: CNN dice que el 51% de los votantes quieren que el Senado condene a Trump por juicio político. Quiero emitir mi voto porque nunca me piden mi opinión. Veamos cómo piensan las personas en Twitter vamos a votar....

- Uno de los *tweets* con 2275 *retweets*, constituye un apoyo abierto al presidente, figura utilizando el slogan *Keep America great again correspondiente* a la campaña de Trump para elecciones de 2020 cuando publica: "El pueblo americano no quiere el *impeachment*, quieren a Donald Trump para mantener a América genial de nuevo".
- Por otro lado, a manera de parodia sobre el juicio político se publicaron *tweets* que expresaban que el evento de *impeachment* iba de adquirir lapiceros que Pelosi colocaba para los presentes. Es importante destacar que los *tweets* publicados con este fin también contenían el hashtag #meme. 3) Con respecto al uso de imágenes y videos se observa una tendencia al *retweet* desde publicaciones oficiales como la de los medios de comunicación y otras instituciones mediáticas gubernamentales, aunque el uso de la imagen y los videomemes fue participativo en el grupo desde un punto de vista risible y burlesco.

#### 4. Conclusiones

Cuando se trata del análisis de los sentimientos es pertinente tener en cuenta el sentido, el significado, el estado de opiniones en los públicos y los códigos en el discurso, los resultados obtenidos en base a las producciones simbólicas son propios del contexto y tiempo de su producción. El presente informe expone conclusiones generales asociadas al tratamiento del impeachment por la esfera pública de Twitter en Estados Unidos. En el caso de estudio la mayor parte del contenido se dio a través de publicación de tweets originales, aunque en un segundo lugar se produjo la conversación (respuestas y menciones respectivamente), y en un menor porcentaje distribución (retweets).

La polaridad y las emociones que mayormente se utilizan en el discurso generado sobre el *impeachment* son neutrales. Los usuarios de la comunidad virtual se dedicaron a distribuir las noticias de los medios de comunicación e instituciones gubernamentales, sobre todo de manera indirecta, con respecto al estado de la cuestión. Sin embargo, los valores afectivos más representados, denotan una duda en la efectividad del juicio para la destitución del presidente; a tal cuestión responden los encuadres de un intento demócrata fallido, las mayoritarias muestras de contenido compartido sobre la ausencia de testigos, documentos y pruebas fácticas en el proceso y la mayoría republicana del Senado.

Así mismo se destacaron usuarios influyentes sobre el total como @RepValDemings, @Aikens\_Josh y @NRSC quienes desde distintas posturas políticas asumen la inseguridad en la condena de Donald Trump en el juicio político ante las circunstancias, argumento más difundido a través de *retweets* por los miembros de la comunidad virtual. En la minería de opiniones comprobamos una relación de los discursos con la cercanía a las elecciones de Estados Unidos en 2020, además de la alusión a la división política entre partidos. Se manifestó una generación de contenido visual que incluía memes para parodiar la situación y el uso de archivos multimediales se incluyó en el 42 % del total de tweets sobre todo para dar sentido o demostrar las aseveraciones del tweet escrito. Estudiar las emociones de los usuarios sobre el #impeachment implica una visión parcial, pues no constituye el total del foro público, pero sí una muestra activa que genera contenidos disponibles para auditar.

#### 5. Referencias bibliográficas

- [1] Abascal, R.; López, E. y Zepeda, S. (2014). Detección de patrones y grupos de sentimientos a partir del análisis de tuits políticos. *Pistas Educativas*, 35(108), 1672-1691. <https://bit.ly/2VWUAtN>
- [2] Aguirre, J.L.; Hernández, C.C; Briceño y Marín, A.M. (2018). Emociones e interacción ciudadana: análisis de la comunicación emocional de los presidentiables en Twitter y Facebook. En B. Gómez y J.L. López (Eds.), *Agenda sociodigital de la campaña presidencial de 2018* (pp. 108-130). Torreón, Coahuila: Universidad Autónoma de Coahuila. <https://bit.ly/2KUaleO>
- [3] Almgren, S.M. & Olsson, T. (2015). Let's get them involved... to some extent: Analyzing online news participation. *Social media + society*, 1(2). <http://doi.org/f3m9j4>
- [4] Arcila, C.; Ortega, F.; Jiménez, J. y Trullenque, S. (2017). Análisis supervisado de sentimientos políticos en español: clasificación en tiempo real de tweets basada en aprendizaje automático. *El profesional de la información*, 26(5), 973-982. <https://bit.ly/2AR9N7H>
- [5] Arnáiz, S.M. y Filardo, L. (2020). Llamamientos feministas en Twitter: ideología, identidad colectiva y reenmarcado de símbolos en la huelga del 8M y la manifestación contra la sentencia de 'La manada'. *Dígitos*, 1(6), 55-78. <http://doi.org/dx72>
- [6] Baeza, R. (2009). Tendencias en minería de datos en la web. *El profesional de la información*, 18(1), 5-10. <http://doi.org/fksfc2>
- [7] Barthes, R. (1971). *Elementos de semiología*. Madrid: Comunicación.
- [8] Baviera, T. (2017). Técnicas para el análisis de sentimiento en Twitter: aprendizaje automático Supervisado y SentiStrength. *Dígitos*, 1(3), 33-50. <https://bit.ly/3dTVqOK>
- [9] BBC News Mundo (31/10/2019). *Por qué Twitter prohibió toda la publicidad política en su plataforma*. <https://bbc.in/2werCeC>
- [10] Briones, G.E. (2018). Nociones de libertad de expresión en disputa: la opinión pública publicada en la prensa ecuatoriana. *Question*, 1(59). <http://doi.org/dx73>

- [11] Burns, A. (2011). Gatekeeping, gatewatching. Real time feed-back: new challenges for journalism. *Brazilian Journalism Research*, 7(11), 117-136. <http://doi.org/gdxvd9>
- [12] Burns, A. & Highfield, T. (2015). From News Blogs to News on Twitter: Gatewatching and Collaborative News Curation. In S. Coleman & D. Freelon (Eds.), *Handbook of Digital Politics*. Edward Elgar Publishing. <http://doi.org/f9sb9z>
- [13] Cansino C.; Calles, J.; Echeverría, M.; Rosanvallon, P.; Pasquino, G.; Morris, I.; Sánchez, J.; Aguilar, E.; Caldevilla, D.; Arditi, B.; Calvillo, J.; Ahuactzin, C. E.; Reguillo, R.; Robles, J.; Torres, L.C.; Meyer, J.A. e Islas, O. (2016). *Del Homo Videns al Homo Twitter: democracia y redes sociales*. México: Benemérita Universidad Autónoma de Puebla.
- [14] Chakrabarti D.; Papadimitriou S.; Modha, D. & Faloutsos, C. (2004). Fully automatic cross-associations. *ACM SIGKDD*, 79-88. <https://bit.ly/2WjowiJ>
- [15] Coddington, M. (2015). Clarifying journalism's quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3), 331-348. <http://doi.org/gc5t4j>
- [16] Cohen, S. & Hamilton, J.T. (2011). Computational journalism. *Communications of the ACM*, 54(10), 66-71. <http://doi.org/cn4t4z>
- [17] Colle, R. (2002). *Explotar la información noticiosa - "Data Mining" en el Periodismo*. España: Universidad Complutense de Madrid.
- [18] Colle, R. (2013). Prensa y "Big Data": El desafío de la acumulación y análisis de datos. *Revista Mediterránea de Comunicación*, 4(1). <http://doi.org/dx76>
- [19] Colle, R. (2017). *Algoritmos, grandes datos e inteligencia en la red. Una visión Crítica*. Colección Mundo Digital de Revista Mediterránea de Comunicación. Universidad de Alicante. <http://doi.org/dx77>
- [20] Conde, M.A.; Pullaguari, K. P. y Prada, O.A. (2019). Comparativa de la competencia mediática en comunicación política vía Twitter durante las campañas electorales de candidatos presidenciales en España, Ecuador y Colombia. *Contratexto*, (32), 41-70. <http://doi.org/dx78>
- [21] Daniel, B. K. (2010) *Handbook of Research on Methods and Techniques for studying virtual communities: Paradigms and Phenomena*. IGI Global. <http://doi.org/cnsjmh>
- [22] De Uribe, C.; Pascual, J.M.; y Gascón, J.F.F (2016). El liderazgo político en la era digital. Un análisis de uso de Twitter por el papa Francisco (@Pontifex). *Obra digital: revista de comunicación*, (11), 85-99. <http://doi.org/dx8v>
- [23] Diakopoulos, N. (2015). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism*, 3(3), 398-415. <http://doi.org/gc5t4g>
- [24] Eco, U. (2000). *Tratado de semiótica general*, Barcelona: Lumen.
- [25] Engesser, S. & Humprecht, E. (2015). Frequency or skillfulness. *Journalism studies*, 16(4), 513-529. <http://doi.org/ggb592>
- [26] Enguix, S. (2017). Impacto político e informativo de las redes sociales: esferas de actuación y comparación con los medios. *Anàlisi: quaderns de comunicació i cultura*, (56), 71-85. <http://doi.org/dgm8>
- [27] Figueira, Á. & Guimarães, N. (2017). Detecting Journalistic Relevance on Social Media: A two-case study using automatic surrogate features. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (pp. 1136-1139). <http://doi.org/dx79>
- [28] Fornari, J.; Abeille, C.; Ferrero, M.; Pérez, G. y Boglione, S. (2019). El potencial de la minería de opinión para la gestión de la reputación online corporativa. *Brazilian Journal of Development*, 5(10), 22310-22320. <http://doi.org/dx8b>
- [29] Foucault, M. (1973). *El orden del discurso*. Buenos Aires, Argentina: Tusquets editores.
- [30] García, J.G.; Henríquez, P. y Herrera, J. (2019). Analítica de Twitter para el estudio de las emociones primarias durante el terremoto de México 2017. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E19), 479-492. <https://bit.ly/3fc1z9c>

- [31] Gomes, F. (2013). *El uso periodístico de las redes sociales por los cibermedios: análisis de perfiles y estudios de caso*. Universidad Rey Juan Carlos, Móstoles, España. <https://bit.ly/3cRZOFy>
- [32] Gómez, M.E. (2014). *Manual imperfecto para políticos en tiempo irreal*. Universidad de La Habana, Cuba.
- [33] González, J.A; Hurtado, Ll. F. y Pla, F. (2018). ELiRF-UPV en TASS 2018: Análisis de Sentimientos en Twitter basado en Aprendizaje Profundo. *Proceedings of TASS*, 37-44.
- [34] Gruzd, A.; Wellman, B. & Takhteyev, Y. (2011). Imagining Twitter as an imagined community. *American Behavioral Scientist*, 55(10), 1294-1318. <http://doi.org/dhbc34>
- [35] Hermida, A. & Young, M. L. (2017). Finding the data unicorn: A hierarchy of hybridity in data and computational journalism. *Digital Journalism*, 5(2), 159-176. <http://doi.org/dx8c>
- [36] Hernández, D. I. (2017). *Irony and sarcasm detection in Twitter: The role of affective content*. Universidad de Valencia, España. <https://bit.ly/2MO5Mno>
- [37] Hootsuite y We are Social. (2020). *Informe 2020 sobre uso de redes sociales*. <https://bit.ly/32FJKKI>
- [38] Iedema, R. (2003). Multimodality, Resemiotization: Extending the Analysis of Discourse as a Multisemiotic Practice. *Visual Communication*, 2(1), 29-57. <http://doi.org/cgtwwk>
- [39] Kwabla, S.; Kwame, N. & Katsriku, F. (2017). Sentiment Analysis of Twitter Feeds using Machine Learning, Effect of Feature Hash Bit Size. *Communications on Applied Electronics (CAE)*, 6(9), 16-21. <http://doi.org/dx8d>
- [40] Leavitt, A.; Burchard, E.; Fisher, D. & Gilbert, S. (2009). The Influentials: New Approaches for Analyzing Influence on Twitter. *Web ecology Project*, 4(2), 1-18. <https://bit.ly/3c4zlxq>
- [41] Lotman, I. (1996). *La semiosfera*. Madrid: Ediciones Cátedra.
- [42] Mancera, A. y Pano, A. (2013a). *El discurso político en Twitter*. Madrid: Editorial Anthropos.
- [43] Mancera, A. y Pano, A. (2013 b). Nuevas dinámicas discursivas en la comunicación política en Twitter. *Círculo de Lingüística Aplicada a la Comunicación*, 56, 53-80. <http://doi.org/dx8f>
- [44] Marín, A.F y Quintero, J.M. (2018). Confianza en el proceso de paz en Colombia en Twitter. *Revista Mexicana de Sociología*, 80(1), 115-137. <https://bit.ly/2zpz3Se>
- [45] Martínez, F. (2013). *Los nuevos medios y el periodismo de medios sociales. Análisis y aplicación de las redes sociales en los medios digitales españoles*. Universidad Complutense de Madrid, España. <https://bit.ly/2MLtArN>
- [46] Martínez, J.J. (2017). *Minería de opiniones mediante análisis de sentimientos y extracción de conceptos en Twitter*. Universidad Complutense de Madrid, España. <https://bit.ly/2UtlWqç>
- [47] Moya, M. y Herrera, S. (2015) Cómo puede contribuir Twitter a una comunicación política más avanzada. *Arbor*, 191(774). <http://doi.org/dx8g>
- [48] O'Halloran, K.L (2012). Análisis del discurso multimodal. *ALED*, 12(1), 75- 97. <http://doi.org/dx8h>
- [49] Orihuela, J. L. (23/06/2011). Los trend topics se han convertido en una agenda social. *La Vanguardia*. (C. Buades, Entrevistador). <https://bit.ly/2PCWso2>
- [50] Ortega, A.; Gutiérrez, E. y Montoyo, A. (2013). SSA-UO: Unsupervised Twitter Sentiment Analysis Second Joint Conference on Lexical and Computational Semantics (\*SEM). *Seventh International Workshop on Semantic Evaluation*, (2), 501-507. <https://bit.ly/2SrWHUw>
- [51] Oviedo, E.A.; Oviedo A. I. y Velez, G.L. (2016). Minería multimedia: hacia la construcción de una metodología y una herramienta de analítica de datos no estructurado. *Revista Ingenierías*, 16(31), 125-142. <http://doi.org/dx8j>
- [52] Pérez, E.; Haber, Y. y Duvergél, M. (2016). Herramientas para el estudio del Nuevo periodismo electrónico o de medios sociales. La perspectiva de framing y análisis crítico discursivo de las etiquetas #The5Free, #Los5Libres y #TodosSomosAmericanos. *Estudios sobre el mensaje periodístico*, 22(2). <http://doi.org/dx8k>

- [53] Pérez, E.; Haber, Y.; Díaz, J. y Zamora, L. (2017). Un modelo periodístico. Interrelación teórica del Nuevo periodismo electrónico y la plataforma Twitter, en el escenario mediático latinoamericano. *Perspectivas de la Comunicación*, 10(1), 157-186. <https://bit.ly/3aYBlyW>
- [54] Pierce, C. S. (1974). *La ciencia de la semiótica*. Buenos Aires: Nueva Visión.
- [55] Portillo, M. (2004). *Culturas juveniles y cultura política: la construcción de la opinión pública de los jóvenes de la Ciudad de México*. Universidad Autónoma de Barcelona, España. <https://bit.ly/2YoKjqj>
- [56] Ramírez, A.C. (2017). *Polaridad en redes sociales y principales indicadores financieros del sector asegurador*. Escuela Colombiana de Ingeniería Julio Garavito, Bogotá, Colombia. <https://bit.ly/3hcrEXj>
- [57] Reyes, A. (2012). *Linguistic-based Patterns for Figurative language Processing: The Case of Humor Recognition and Irony Detection*. Universidad de Valencia, España. <https://bit.ly/3dRDZ1c>
- [58] Reyes, J.A.; Paniagua, F. y Sánchez, L. (2017). Minería de opiniones centrada en tópicos usando textos cortos en español. *Research in Computing Science*, 134, 151-162. <http://doi.org/dx8m>
- [59] Rheingold, H. (1996). *La comunidad virtual*. Barcelona: Gedisa.
- [60] Rodríguez, K. y Haber, Y. (2017). La influencia social de los medios de comunicación en Twitter. *Enunciación*, 22(1), 97-108. <http://doi.org/dx8n>
- [61] Saussure, F. (1961). *Curso de Lingüística General*. Buenos Aires: Losada.
- [62] Scolari, C. A. (2018). *Las leyes de la interfaz: Diseño, ecología, evolución, tecnología*. Barcelona: Gedisa.
- [63] Suau, G.; Percastre, S.J.; Palá, G. y Pont, C. (2017). Análisis de la comunicación de emergencias en Twitter. El caso del ébola en España. En *Uso y aplicación de las redes sociales en el mundo audiovisual y publicitaria* (pp. 119-130). McGrawhill Education. <https://bit.ly/2SvpP1n>
- [64] Vallejo, A. M. (2018). Comunicar emociones en el discurso metapolítico de Twitter: el caso de #Maduro versus @NICOLASMADURO. *Observatorio (OBS)*, 12(3), 175-194. <http://doi.org/dx8p>
- [65] Vález, M., y Codina, L. (2018). Periodismo computacional: evolución, casos y herramientas. *El profesional de la información (EPI)*, 27(4), 759-768. <http://doi.org/dx8q>
- [66] Van Hee, C. (2017) *Can machines sense irony?* Universidad de Gante, Bélgica.
- [67] Vásquez, L.C. (2019). Semiosis entre palabras e imágenes en Twitter. ¿Cómo construyen significado los candidatos presidenciales de las elecciones 2017 de Chile? *Galaxia*, 42, 40-56. <http://doi.org/dx8r>
- [68] Ventura, A.S. (2016). Argumentación y discurso político en Twitter. Análisis de la campaña presidencial argentina 2011. *Aposta: Revista de Ciencias Sociales*, 69, 39-66. <https://bit.ly/2xwQ4Je>
- [69] Verbeke, M.; Berendt, B.; D'Haenens, L. & Opgenhaffen, M. (2017). Critical news reading with Twitter? Exploring data-mining practices and their impact on societal discourse. *Communications*, 42(2), 127-149. <http://doi.org/dx8s>
- [70] Verón, E. (1987). *La semiosis social*. Buenos Aires: Gedisa.
- [71] Vilares, D.; Alonso, M.A. y Gómez, C. (2013). Una aproximación supervisada para la minería de opiniones sobre tweets en español en base a conocimiento lingüístico. *Procesamiento del Lenguaje natural*, 51, 127-134. <https://bit.ly/3c1Ef3Y>
- [72] Vilariño, D.; Zepeda, C.; Sanzón, Y.M.; Carballido, J.L.; Medina, C. y Flores, G. (2015) Aplicación del modelo Léxico-Sintáctico para detectar la polaridad de opiniones sobre profesores. *Pistas Educativas*, 36(113), 228-239. <https://bit.ly/2KUtJs1>
- [73] Welbers, K. y Opgenhaffen, M. (2019). News through a social media filter. Different perspectives on immigration in news on website and social media formats. *Journalism and Translation in the Era of Convergence*, 85-105. <http://doi.org/dx8t>
- [74] Wodak, R. y Meyer, M. (2003). *Método de análisis crítico del discurso*. Barcelona: Gedisa.
- [75] Wu, S.; Hofman, J.; Mason, W. & Watts, D. (2011). Who Says What to Whom on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (pp. 705-714). <http://doi.org/cxcjbn>

## Notas

1. Comunidades virtuales: Grupos sociales formados por intermedio de los ordenadores (Rheingold, 1996: 15). Las *comunidades virtuales* son agregados sociales que surgen de la Red cuando una cantidad suficiente de gente lleva a cabo estas discusiones públicas durante un tiempo suficiente, con suficientes sentimientos humanos como para formar redes de relaciones personales en el espacio cibernético.

Las comunidades virtuales son el resultado de la sociabilidad natural característica de los seres humanos, que cuando encuentran el sitio adecuado para la creación de vínculos a través de Internet, utilizan la comunicación mediada por ordenador para llevar a cabo agregados sociales cohesionados por una relación de confianza basada en intereses comunes, sentimiento de identificación con los demás e intercambio constante de experiencias (Gomes, 2013).

2. Datos obtenidos a través del informe de 2020 ofrecido por las plataformas Hootsuite y We are social sobre el uso de las redes sociales a nivel global. Disponible en: <https://bit.ly/32FJKKI>

3. BBC News Mundo. (31 de octubre de 2019). Por qué Twitter prohibió toda la publicidad política en su plataforma. Disponible en: <https://bbc.in/2werCeC>

4. Big Data: Conjunto de datos o combinaciones de conjunto de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad), dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como base de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles. Su naturaleza compleja parte de la gran cantidad de datos no estructurados generados por las tecnologías modernas como las redes sociales. Lo que hace útil al análisis de big data es que proporciona puntos de referencia en la recopilación de grandes cantidades de datos que son filtrados y calificados de manera rápida y eficiente.

5. Algoritmo: Un algoritmo (del griego y latín, *dixit algorithmus* y este a su vez del matemático persa Al-Juarismi) es un conjunto prescrito de instrucciones o reglas bien definidas, ordenadas y finitas que permite llevar a cabo una actividad mediante pasos sucesivos que no generen dudas a quien deba hacer dicha actividad. Dados un estado inicial y una entrada, siguiendo los pasos sucesivos se llega a un estado final y se obtiene una solución (Colle, 2017: 6). Normalmente un programa de minería de datos utiliza varios algoritmos para filtrar información y dar resultados sobre una muestra dada.

6. Clustering: es una técnica de minería de datos que identifica de forma automática agrupaciones de acuerdo a una similitud entre ellos (concordancia). Se habla de búsqueda de patrones o técnicas de clustering.

7. Polaridad: es un valor que se le asigna a un término que expresa opinión dependiendo del significado lingüístico de la palabra y se basa en metodologías de análisis de los sentimientos o minería de opiniones. El valor de la polaridad puede variar entre diferentes rangos entre ellos negativo, neutro y positivo. (Ramírez, 2017: 47)

8. Un grafo es una unidad grupal que relaciona nodos conectados a través de aristas. En el caso particular se refiere a las relaciones sociales que se dan entre los usuarios de la comunidad virtual y sus vías de enlace.

9. El análisis de la representación simbólica se fundamenta en la interpretación de elementos que generan el sentido. La manifestación de este sentido se circunscribe a una dimensión de significación asociada a una dimensión sociocultural. Véase Saussure, (1961); Barthes, (1971); Pierce, (1974); Verón, (1987); Lotman, (1996); Eco, (2000).

10. La influencia social se valora como las reacciones de *retweet*, me gusta y respuesta de cierto contenido o usuario sobre otros del grupo. Es decir, la potencialidad de la acción de un usuario (por ej.: *tweet*) de iniciar futuras acciones (por ej.: *retweet* o respuesta) de otros en la red, lo cual se evidencia a partir de la acción de crear un mensaje (*tweet*) y suscitar su posterior re-difusión (*retweet*) por su red de contactos y de las redes de quienes lo reciben y deciden reenviarlo (Leavitt, Burchard, Fisher y Gilbert, 2009: 5). En Twitter persisten dos tipos de influencia: una basada en la conversación y otra basada en el contenido.

