



Análisis de texto en Twitter para las elecciones a la Alcaldía de Bogotá 2019

DOI: <https://doi.org/10.21158/21451494.v8.n0.2017.2700>

MarcelaGarcía-Guerrero
Universidad de La Salle
marcelagarcia61@unisalle.edu.co

Laura Sofía Aguirre-Peña
Universidad de La Salle
laguirre35@unisalle.edu.co



1 Estudiante del programa de Negocios y Relaciones Internacionales de la Universidad de La Salle- Colombia.
ORCID: <https://orcid.org/0000-0001-9687-3348>

2 Estudiante del programa de Negocios y Relaciones Internacionales de la Universidad de La Salle- Colombia.
ORCID: <https://orcid.org/0000-0002-4421-2158>

Cómo citar este artículo:

García-Guerrero, M.; Aguirre-Peña, L. S. (2017). Análisis de texto en Twitter para las elecciones a la Alcaldía de Bogotá 2019. Revista Comunicación, cultura y política, 8, 22-48. DOI: <https://doi.org/10.21158/21451494.v8.n0.2017.2700>

Resumen

Una de las técnicas en el campo de la big data que se utiliza para indagar la relación de las personas con respecto a algún tema en específico es la minería de texto, que consiste en utilizar palabras o textos con el propósito de inferir análisis robustos de una situación. En el presente artículo se presenta un análisis de texto de las interacciones en Twitter por parte de los candidatos y seguidores en los últimos cuatro días de campaña a la Alcaldía de Bogotá 2019, con el fin de analizar el nivel de aceptación que tienen los seguidores frente a las publicaciones de los candidatos en torno a sus campañas en este periodo previo al ejercicio de votación. La técnica utilizada fue análisis de texto, para lo cual se extrajeron los trinos de los seguidores de los candidatos a la Alcaldía de Bogotá mediante el uso del software estadístico R-Studio y API. El resultado muestra cuatro gráficos correspondientes a la frecuencia de palabras de los seguidores de cada candidato. Se inicia con el candidato Miguel Uribe, seguido por el candidato Carlos Galán, luego el candidato Hollman Morris y, finalmente, la candidata Claudia López, de manera que se muestra cómo entre más diversidad de colores y el aumento del tamaño de las palabras mayor es la frecuencia en el uso de estas, lo que genera una mayor comunicación entre el candidato y los seguidores a fin de promover las campañas.

Palabras clave: big data; minería de texto; análisis de texto; redes sociales; Twitter; elecciones Alcaldía de Bogotá.

Analysis of Twitter text for Bogotá mayoral elections 2019

Abstract

One of the techniques that is used to investigate the relationship of people with respect to a specific topic in the field of big data is text mining, which is basically using words or texts with the purpose of inferring robust analysis of a situation. This article presents a text analysis of the interactions on Twitter among candidates and followers during the last four days of Bogotá mayoral campaign 2019, with the purpose of analyzing the level of acceptance that followers have of the publications of the candidates regarding their campaigns during this period prior to the voting exercise. We used the text analysis technique, for which the trills of the followers of the candidates for Mayor of Bogotá were extracted using the statistical software R-Studio and API. The results show four graphs corresponding to the word frequency of the followers of each candidate. It begins with candidate Miguel Uribe, followed by candidate Carlos Galán, then candidate Hollman Morris, and finally candidate Claudia López, showing that the more diverse the colors and the larger the word size, the more frequent the use of words, which generates greater communication between the candidate and the supporters to promote the campaigns.

Keywords: *big data; text mining; text analysis; social networks; Twitter; Bogotá mayoral elections.*

Análise de textos do Twitter para as eleições à prefeitura de Bogotá 2019

Resumo

Uma das técnicas da área de big data que se utiliza para investigar a relação das pessoas a respeito de algum tema específico é a mineração de texto, que consiste em utilizar palavras ou textos com o objetivo de inferir análises robustas de uma situação. Este artigo apresenta uma análise textual das interações no Twitter dos candidatos e seguidores nos últimos quatro dias da campanha à prefeitura de Bogotá em 2019, com o objetivo de analisar o nível de aceitação que os seguidores têm diante de às publicações dos candidatos relativas às suas campanhas antes das votações. A técnica utilizada foi a análise de textos, em que foram extraídos os trigramas dos seguidores dos candidatos à prefeitura de Bogotá por meio dos softwares estatísticos R-Studio e API. O resultado mostra quatro gráficos correspondentes à frequência de palavras dos seguidores de cada candidato. Inicia com o candidato Miguel Uribe, seguido pelo candidato Carlos Galán, depois o candidato Hollman Morris e, por último, a candidata Claudia López, de forma que quanto maior a diversidade de cores e do aumento no tamanho das palavras, maior é a frequência na utilização destas, o que gera maior comunicação entre o candidato e os seguidores para divulgação das campanhas.

Palavras-chave: *big data; mineração de texto; análise de texto; redes sociais; Twitter; eleições para à prefeitura de Bogotá.*

Analyse des messages Twitter des campagnes électorales à la mairie de Bogotá de 2019

Résumé

L'une des techniques du domaine des métadonnées utilisée pour étudier les relations interpersonnelles sur un sujet spécifique est l'exploration de texte qui consiste à utiliser des mots ou des textes dans le but de déduire des analyses robustes d'une situation donnée. Cet article présente une analyse textuelle des interactions des candidats et leurs adhérents sur Twitter durant les quatre derniers jours de la campagne pour la mairie de Bogotá de 2019, et analyse le niveau d'acceptation des adhérents des publications des candidats en campagne. La technique utilisée est l'analyse des textes envoyés par les adhérents, extraits à l'aide des logiciels statistiques R-Studio et API. Les résultats montrent quatre graphiques correspondant à la fréquence des mots utilisés par les adhérents de chaque candidat. L'article analyse d'abord les résultats obtenus par le candidat Miguel Uribe, suivi du candidat Carlos Galán, du candidat Hollman Morris et, enfin, de la candidate Claudia López. L'analyse révèle la diversité des couleurs et l'augmentation de la taille des mots et leur fréquence d'utilisation qui génèrent une meilleure communication entre le candidat et ses adhérents durant la campagne électorale.

Mots clés: big data; analyse de texte; réseaux sociaux; Twitter; élection à la mairie de Bogotá.

1. Introducción

En los últimos años el auge de las redes sociales en el ámbito de la comunicación política ha sido determinante para generar cambios trascendentales y fortalecer la participación crítica en este campo: «*One of the fundamental ways the advent of social media services has changed political communication is that social media services have become additional channels in the political communication space*» (Jungerr, 2014, p. 4).

Para empezar, un primer momento en el que se evidenció este vínculo, de acuerdo con Segado-Boj, Díaz-Campo y Lloves-Sobrado (2015), fue «*especially after the milestone set by Barack Obama's 2008 election campaign*» (p. 157). En este caso las redes sociales influenciaron de manera positiva al candidato. En Colombia se han evidenciado pocos estudios que vinculan la política y las redes sociales. Según Gómez (2016), «se han realizado pocos estudios de comunicación política que se dediquen específicamente a las elecciones y el uso de redes sociales» (p. 19).

En este sentido, el objetivo general del presente artículo es analizar las unidades de texto de las que se infiere el nivel de aceptación que tienen los seguidores frente a las publicaciones de los candidatos en torno a sus campañas en los últimos cuatro días antes de las elecciones. Para esto se indagaron los diferentes estudios previos que existen sobre el tema, con el fin de conocer vacíos. Se utilizó un *software* estadístico —R-Studio— para la extracción de datos en tiempo real sobre información de candidatos-seguidores.

Este artículo de investigación se divide en cinco secciones. La primera es esta introducción, posteriormente se expone el perfil y la dinámica de cada uno de los candidatos en Twitter durante el periodo previo a las elecciones. La tercera sección detalla la metodología empleada y el uso de herramientas estadísticas, la cuarta presenta el análisis de resultados y, por último, se da paso a las conclusiones.

Es importante resaltar que en el desarrollo del presente artículo de investigación, en la revisión de literatura para el periodo 2019-2020, la información encontrada fue muy limitada, ya que son pocos los trabajos de investigación en Colombia que abordan esta temática.

1.2 La dinámica en Twitter de los candidatos a la Alcaldía de Bogotá-2019

El 27 de julio se realizó la inscripción de los candidatos oficiales para las elecciones a la Alcaldía de Bogotá, de manera que se generaron dos grandes bloques: los alternativos, representados por Claudia López, y la derecha, representada por Miguel Uribe Turbay. Así mismo, con un movimiento independiente, Carlos Fernando Galán, y representando a Colombia Humana se postuló Hollman Morris (Pérez, 2019).

Los candidatos mencionados emprendieron una dinámica de campaña a través de las redes sociales, principalmente en Twitter, en la que es posible identificar el perfil de los candidatos, su impacto en determinados grupos sociales y su modo de comunicación y participación entre candidatos-seguidores. Para empezar, el candidato por el movimiento Avancemos, Miguel Uribe Turbay, presenta un comportamiento controversial en Twitter. De acuerdo con Semana (El repunte de Hollman Morris, 8 de junio de 2019), «Miguel Uribe Turbay también ha logrado altísimos números de interacciones y alcance gracias a sus constantes controversias y ataques a Claudia López, aunque ella todavía le lleva mucha ventaja» (párr. 3).

Con respecto al manejo de las redes sociales y las comunicaciones, según Pinilla (2019), «la estrategia de comunicación está basada en enviar el mensaje de que Uribe representa la continuidad del proyecto de ciudad de Peñalosa y de que hay que construir sobre lo construido». Además, el candidato durante su campaña invirtió COP 1667 millones que corresponden a medios de comunicación y estrategia de publicidad.

Por otra parte, el candidato Luis Fernando Galán, de acuerdo con Semana (El repunte de Hollman Morris, 8 de junio de 2019), «tuvo una importante exposición en medios, los comentarios en su mayoría eran positivos, pero no generó mayor conversación» (párr. 1). Se evidencia que tiene bajo reconocimiento comparado con los dos candidatos que presentan mayores interacciones como Miguel Uribe y Claudia López.

Sin embargo, el uso de retuits por parte de ciudadanos que le brindan apoyo a su campaña ha sido un punto a destacar en las redes como estrategia política. Además, la jefa de Comunicaciones es Margarita Castro, creadora de una agencia de comunicación pública y medios digitales, por lo que en su equipo tienen amplia experiencia en medios de comunicación masivos (Pinilla, 2019). La cuenta de Hollman Morris presenta trinos relacionados con el panorama político del país y algunos sobre Bogotá; sus apariciones en la red social presentan similitud a su aliado político Gustavo Petro. Así mismo, es reconocido de forma negativa debido a polémicas de su vida personal.

Cabe resaltar que durante su precampaña no definió un hashtag que lo hiciera conocido en redes; sin embargo, en la última etapa de campaña definió uno que tuvo una dinámica mínima. El candidato tiene como representante para el manejo de redes sociales a María Paula Fonseca, periodista y productora audiovisual de la Universidad Pontificia Javeriana y candidata a Magíster en Ciencia Política de la Universidad de los Andes (Pinilla, 2019).

Para finalizar, la candidata Claudia López es una de las candidatas más activas en Twitter, si se tiene en cuenta que es la candidata que más seguidores tiene, lo que la convierte en un personaje popular y presente en la mayoría de debates. Su campaña cuenta con Daniel Winograd, publicista encargado de los contenidos de comunicaciones, también la acompaña el comité que prepara los debates, compuesto por Augusto Reyes, Ana María Ruiz y Guillermo Cuéllar, quienes coordinan temas tales como entrevistas, redes sociales y publicidad (Pinilla, 2019).

2. Metodología

La metodología de la investigación se desarrolló en tres fases. En la primera, se determinan las cuentas de Twitter de los cuatro candidatos a la Alcaldía en la última etapa de campaña. Estas son: @ClaudiaLopez, @MiguelUribeT, @CarlosFGalan y @HOLLMANMORRIS. Por medio de estas cuentas se observa el comportamiento de los seguidores al indagar la interacción con el tipo de *hashtag* que utilizan en la campaña y el uso de emoticones, entre otros aspectos de su comunicación.

En la segunda fase, mediante el programa R-studio y a través de la API de Twitter se inicia la recolección de información. De este modo, se filtra la información de las cuentas por fecha, se procede a la creación de un *data frame* para almacenar los datos de los trinos emitidos por los seguidores y candidatos, luego se limpia el texto generado, es decir, se suprimen los puntos, las comas y los símbolos de cada trino. Posteriormente, mediante la función *Corpus*, se tokenizan los trinos de modo que se conviertan en una unidad de texto sencilla.

A continuación, mediante la función *Stopwords*, eliminamos las palabras más repetidas con el fin de limpiar la base de datos que se genera; de esta forma el análisis de texto tiende a ser más preciso. Luego se genera una nueva matriz con los datos limpios y mediante la función *Inspect* obtenemos las palabras más frecuentes. Por último, se crea una matriz con los trinos limpios que contengan las palabras más relevantes.

La última fase consiste en generar una relación de frecuencia mediante una nube de palabras que refleje las palabras más usadas por los seguidores y candidatos durante los últimos cuatro días de campaña; así, entonces, se destaca el principal tema de interés durante cada día. Lo anterior con el fin de obtener una aproximación más acertada en la relación candidato-ciudadano e identificar los intereses de los ciudadanos, para lo cual esta red social ha sido determinante al generar un espacio libre de participación política y afianzar la comunicación entre ambas partes.

3. Análisis de resultados

El análisis de texto realizado muestra los resultados obtenidos durante los cuatro días previos a las votaciones para la Alcaldía de Bogotá del 2019. Dada la naturaleza de la investigación, se carece de un modelo estadístico para tratar los datos obtenidos, no obstante, se emplea un software estadístico —R-studio—, el cual permite realizar una técnica mediante la API de Twitter a fin de recopilar información de diferentes cuentas oficiales, de manera que es posible obtener acceso directo a la información y ejecutar el debido análisis de texto.

Con el fin de proceder, por cada candidato se exponen cuatro gráficos por día, correspondientes a la frecuencia de palabras de los seguidores. Se inicia con el candidato Miguel Uribe, se continúa con el candidato Carlos Galán, enseguida con el candidato Hollman Morris y se finaliza con la candidata Claudia López.

Figura 1.
Frecuencia de palabras del candidato Miguel Uribe – 21 de octubre del 2019

Fuente. Elaboración propia.



3.1 Candidato I: Miguel Uribe Turbay

En la figura 1 se observa el comportamiento de los seguidores de Miguel Uribe Turbay el 21 de octubre del 2019. Se evidencia que el nombre del candidato es el eje central. La palabra que más captura la atención es «avancemos»; es la principal de la campaña y genera identificación en la forma de comunicación en su campaña virtual.

En el segundo día de análisis, el 22 de octubre del 2019, la nube de palabras (véase la Figura 2) mantiene la tendencia en los términos con mayor frecuencia, los cuales son «avancemos», «encuesta», «Bogotá» y «apoyo». Este candidato se centra en exponer términos alusivos a su campaña; hasta ese momento no se observa protagonismo de palabras sobre sus propuestas o problemáticas de la ciudad.

Con respecto al comportamiento de los seguidores, en la nube de palabras tienen protagonismo las siguientes cuentas: @linamcardonag y @totorangelm, las cuales pertenecen a ciudadanas muy activas que



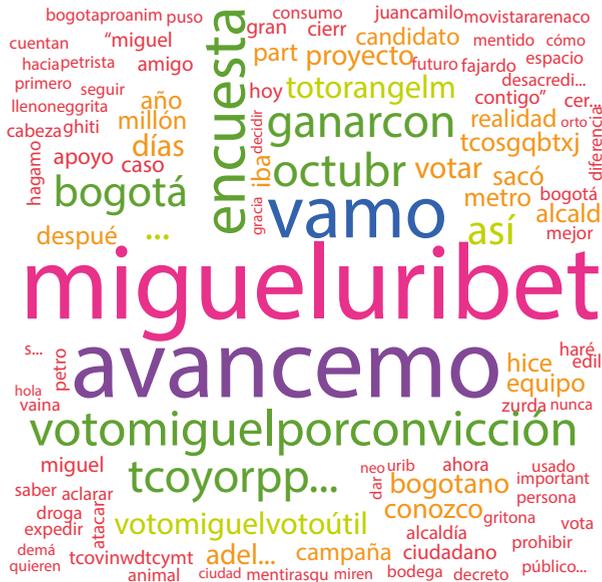
Figura 2.
Frecuencia de palabras del candidato Miguel Uribe – 22 de octubre del 2019

Fuente. Elaboración propia.

promueven la campaña a través de sus cuentas. Además, @JCMontanaParra y @kroarbelaez son candidatos al Concejo por el Centro Democrático y Cambio Radical, respectivamente, que han demostrado afinidad con las propuestas del candidato. Por último, @CARLOSFMEJIA, senador de la República que pertenece al partido Centro Democrático.

La nube de palabras de la figura 3 correspondiente al día 23 de octubre del 2019, resalta el protagonismo de las palabras alusivas a su campaña, tales como «migueluribet», «votomiguelporconviccion» y «alcaldía». Una de las razones es el cierre de campaña, en el que los candidatos tienen mayor dinámica a través su hashtag de campaña; en este caso es #VotoMiguelPorConviccion. Cabe destacar de nuevo a @CARLOSFMEJIA, senador de la República y miembro del partido Centro Democrático, quien ha manifestado su apoyo a la campaña de Miguel Uribe.

Figura 3.
Frecuencia de palabras del candidato Miguel Uribe – 21 de octubre del 2019
Fuente. Elaboración propia.



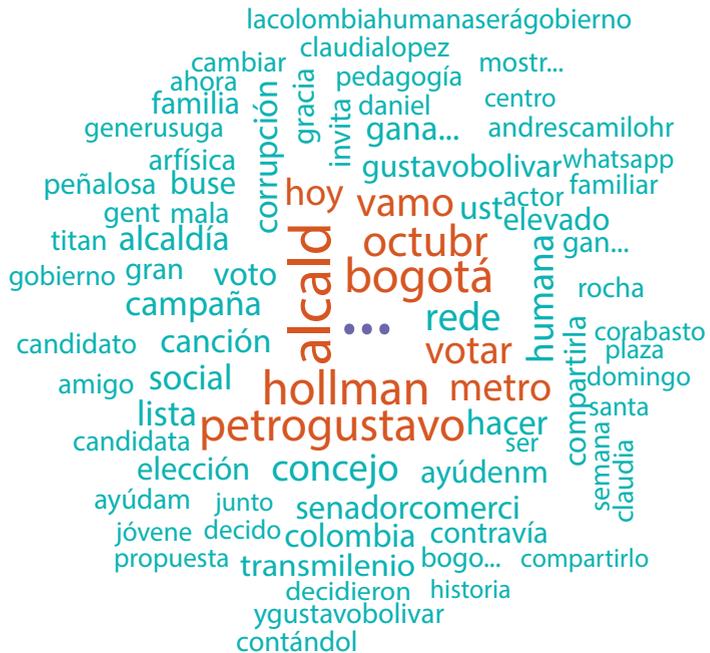
3.3. Candidato III: Hollman Morris

En la figura 9 se evidencia que algunas de las palabras más representativas son «Hollman», «alcalde», «Bogotá», «hoy», «metro» y, evidentemente, «Petro Gustavo». No se evidencia el uso de palabras más variadas, ya que el gráfico solo muestra dos colores y pocas palabras representativas. Esto puede darse debido a que el perfil que maneja el candidato está más asociado a otros medios de comunicación masiva como, por ejemplo, periódicos o TV.

Por otra parte, con respecto a las cuentas que se evidencian en el gráfico @ petrogustavo es la cuenta más destacada, así como @GustavoBolivar, quienes son los seguidores más cercanos a la campaña. También se evidencia la mención de la cuenta de Claudia López, quien es la candidata con quien más ideas comparte.

Figura 9.
Frecuencia de palabras candidato Hollman Morris– 21 de octubre del 2019

Fuente. Elaboración propia.



Entre los seguidores más activos en la campaña se encuentran: @lacarinsaam, @AngelicaLozanoC, @pedroalerta y @rcnradio. Se destacan debido a las constantes noticias emitidas por el cierre de campaña y al presentar los resultados de las encuestas. Por su parte, @maferojas, candidata al Concejo de Bogotá y quien ha brindado un importante apoyo y acompañamiento a Claudia López, también manifiesta en su cuenta el apoyo a la alcaldía de la candidata.

Durante la campaña, Claudia López fue una de las candidatas más activas en Twitter. A través de los gráficos es posible identificar que durante la campaña los términos con mayor frecuencia son alusivos a las propuestas y problemáticas de la ciudad, a diferencia de otros candidatos que se centran en su imagen y campaña. Cabe resaltar que en el análisis se evidencia un cambio permanente en las tendencias y la frecuencia de las palabras más utilizadas por la candidata, lo que genera dinámica en la información.



Figura 16.
Frecuencia de palabras candidata Claudia López– 24 de octubre del 2019

Fuente. Elaboración propia.

4. Conclusiones

Los resultados del trabajo de investigación evidencian la dinámica emprendida por los políticos en redes sociales y la importancia que en la actualidad tiene el manejo de la información para el ejercicio de la política. Sin embargo, el comportamiento de los seguidores de los candidatos puede verse empañado por la manipulación de la información que tiende a afectar el ejercicio de la democracia.

En los gráficos expuestos se identificaron aspectos claves que representan una manipulación en la información, los cuales se describen a continuación. Lo primero es las dimensiones de las ilustraciones, ya que al presentar una simetría casi perfecta se evidencia un manejo especial de la información. Esto puede ser diseñado por expertos a partir de cuentas falsas, conocidos como bots, con el fin de manipular los resultados en las frecuencias en algunos términos y viralizar las cuentas con el objetivo de ganar más seguidores que pueden representar votos.

Segundo, en el manejo de la información se destacan los términos más relevantes alusivos a la campaña, es decir, el eslogan o la imagen del candidato. No obstante, los temas coyunturales de mayor relevancia para el ciudadano evidencian una menor frecuencia, lo cual refleja que en los últimos días de campaña los candidatos pretenden destacar el partido que representan y generar debates políticos entre sus opositores.

Ahora bien, se debe tener en cuenta que el manejo de las redes sociales está en manos de personal especializado, quienes generan una alteración en los datos recopilados y en los resultados finales de la investigación: no obstante, se debe destacar la dinámica con respecto a la participación de seguidores y la generación de debates de Claudia López y Miguel Uribe durante el 21 de octubre hasta el 24 de octubre del 2019.

Adicionalmente, los gráficos expuestos reflejan la empatía que se genera mediante la interacción directa entre candidato-seguidor en tiempo real. Esta característica se logra identificar mediante las palabras que transmiten emociones de los usuarios y seguidores ante las propuestas, lo cual refleja los principales temas de interés de acuerdo con la coyuntura del país. Esta dinámica debe tenerse en cuenta para enfrentar los nuevos retos de hacer política y ejecutar campañas electorales; no obstante, se corre el riesgo de alteraciones en los datos y la información, por las razones expuestas.

Finalmente, en este tipo de estudios se puede realizar un análisis más riguroso con respecto a las emociones y la comunicación de los seguidores-candidatos con el fin de generar un modelo predictivo que aporte mayor seguridad en los procesos electorales y, de este modo, identificar el poder de las redes sociales para persuadir a los usuarios en el propósito de ganar las elecciones. Esto exige un mayor conocimiento en temas de programación debido al tamaño de los datos a trabajar en tiempo real.

Referencias

¿El repunte de Hollman Morris en redes sociales es negativo o positivo? (8 de junio de 2019). *Semana*. Recuperado de <https://bit.ly/35w6CxW>

Gómez, L. (2016). *Comunicación política por Twitter. Colombia: a un tweet de la democracia* (Tesis de doctorado). Universidad Sergio Arboleda, Colombia. Recuperado de <https://bit.ly/3oq6SHz>

Jungherr, A. (2014). The logic of political coverage on Twitter: temporal dynamics and content. *Journal of Communication*, 64(2), 239-259. DOI: <https://doi.org/10.1111/jcom.12087>

Pérez, N. (5 de agosto de 2019). Bogotá: ¿cómo quedaron los candidatos? *Razón Pública*. Recuperado de <https://bit.ly/34sZHGw>

Pinilla, Á. (16 de septiembre de 2019). Los que rodean a los candidatos en Bogotá. *La Silla Vacía*. Recuperado de <https://bit.ly/35CJldO>

Segado-Bog, F.; Díaz-Campo, J.; Lloves-Sobrado, B. (2015): Latin American leader son Twitter. Old uses for new media during political crisis. *Revista Latina de Comunicación Social*, 70, 155 – 173. DOI: <https://doi.org/10.4185/RLCS-2015-1040en>

