
Towards improving data-driven public sector: an approach using Twitter API

Victor Eduardo Romero Cueva

Jimmy Rafael Landaburú Mendoza

Universidad Técnica Estatal de Quevedo

Quevedo, Ecuador

Abstract: *At the present time, the governance and management of public affairs is marked by the immediacy of information. The global reality requires the adoption of digital strategies that allow the collaboration of the group of actors involved: citizens, companies and entities. In this sense, Twitter has become an essential means of communication when addressing the problems of contemporary societies. This microblogging tool has shown its usefulness in decision-making, and for this it has made available an API (Application Programming Interface) that allows the extraction of data of interest. In this research experience, it has been seen that it is possible to use a data-driven approach to contribute to the evaluation of the public agenda. For this, the GADs Municipales of Quito and Guayaquil were selected from which 2 leading agencies in social responsibility, respectively derive. The R statistical software was used to perform data mining, in addition to the subsequent analysis and visualization of trends. The results of this study showed that, although it is true that there are initiatives aimed at the integral development of citizenship, there is a marked contrast with respect to the way in which relationships are built with target audiences. This, added to the low priority towards digital transformation on the public agenda, suggests that local governments are still far from what smart cities are.*

Key words: *Digital governance, data-driven, public agenda, social media, sentiment analysis.*

Hacia la mejora del sector público impulsado por datos: un enfoque utilizando la API de Twitter

Resumen: *En el momento presente, la gobernanza y gestión de los asuntos públicos está marcada por la inmediatez de la información. La realidad global exige la adopción de estrategias digitales que permitan la colaboración del conjunto de actores involucrados: ciudadanos, empresas y entidades. En este sentido, Twitter ha pasado a ser un medio de comunicación imprescindible a la hora de abordar los problemas de las sociedades contemporáneas. Esta herramienta de microblogging ha mostrado su utilidad en la toma de decisiones, y para ello ha puesto a disposición una API (Interfaz de Programación de Aplicaciones) que permite extraer datos de interés. En esta experiencia de investigación se ha visto que es posible emplear un enfoque data-driven para contribuir en la evaluación de la agenda pública. Para ello, se seleccionó los GADs Municipales de Quito y Guayaquil de los que derivan respectivamente 2 agencias referentes en responsabilidad social. Se utilizó el software estadístico R para realizar la minería de datos, además del posterior análisis y visualización de tendencias. Los resultados de este estudio mostraron que, si bien es cierto existen iniciativas orientadas al desarrollo integral de la ciudadanía, hay un contraste marcado respecto a la forma en que se construyen relaciones con las audiencias objetivo. Esto, sumado a la escasa prioridad hacia la transformación digital en la agenda pública, deja entrever que los gobiernos locales distan aún de lo que son las smart cities.*

Palabras clave: *Gobernanza digital, data-driven, agenda pública, social media, análisis de sentimientos*

1. Introducción

No cabe duda de que la pandemia de Covid-19 ha dejado en evidencia las deficiencias estructurales del llamado mundo desarrollado, y ni qué hablar de los territorios emergentes. Pero también ha demostrado que la tecnología juega un rol determinante a la hora de enfrentar colectivamente la crisis en mención. En este sentido, es importante resaltar que la crisis no es solo un problema económico o de salud pública, sino también de transformación digital.

El último informe del (BID, 2020) reporta que solo 17 de 32 países en América poseen políticas públicas de ciberseguridad. Ecuador aparece como una economía que aún está desarrollando su estrategia nacional. Pero, ¿se podía esperar grandes avances del país que recién desde el 2018 se integró a la Alianza para el Gobierno Abierto? Lógicamente no. Además, si el gobierno de turno apenas se ha dado cuenta del crimen organizado que impera en los centros penitenciarios, no es difícil suponer el escenario en materia de cultura de datos o, mejor dicho, data-driven culture.

Ahora bien, ¿es lo mismo data culture que data-driven culture? En su estudio crítico sobre cultura de datos, (Acker & Clement, 2019) lo analizan detenidamente:

Understanding data cultures as underwritten by collections of data (as relata) means understanding data cultures as phenomenon shaped by ideas about the cultivation and production of data that reflect epistemologies about, for example, ordering, classification, and standards. Like an underpainting, data is an initial layer, a starting point against which the authors look at the cultures that have been painted in the background of data collecting and collections.

Una definición similar presentan (Kremser & Brunauer, 2019):

An organizational culture is the collection of common values, behaviors and norms which are postulated by top management and are followed by most individuals within an organization. Data culture therefore refers to the values, behaviors and norms shared by most individuals within an organization with regards to data-related issues.

Al respecto, es innegable que ambas posturas guardan relación con el proceso de adquisición-transmisión de conductas. Lo cual indica que, este tipo de cultura se basa en la idea de una sociedad que genera, e incluso utiliza datos. Esto hace que surjan cuestionamientos serios con respecto a la protección de datos personales, por un lado, pero por el otro, esta digitalización de la sociedad está trayendo más beneficios de lo que se esperaba en anteriores modos de producción.

En cuanto al data-driven culture, (Berndtsson, Forsberg, Stein, & Svahn, 2018) señalan que “it is characterized by a decision process that emphasise testing and experimentation, where data outweighs opinions, and where failure is accepted as long as something is learnt from it”. Es decir, se centra en aprovechar las ventajas derivadas del uso de los datos para la toma de decisiones, mas no en la intuición. Dicho de otra forma, implica pasar del conocimiento a la acción. Por ejemplo, basta mencionar el caso de la economía colaborativa o *sharing economy*, en la que gracias a la inmediatez de la información se provee servicios de calidad a un sinnúmero de usuarios.

No obstante, ocurre a menudo que levantar información de la forma tradicional toma su tiempo, así como también su procesamiento. A esto hay que añadir el costo respectivo, sobre todo, cuando se realiza cara a cara. Y es que, en la era digital, las tendencias han cambiado de forma radical, pues las personas migran continuamente a las plataformas virtuales. Twitter es una de ellas que, más que una red social, es un medio de comunicación. En

palabras de (Sánchez Holgado, Martín Merino, & Blanco Herrero, 2020):

Tres billones de personas alrededor del mundo expresan sus pensamientos y opiniones de modo habitual a través de redes sociales. Twitter es una de esas redes sociales; se trata de un servicio de microblogging que combina características de blog, mensajería instantánea y redes sociales. . . . Los usuarios de Twitter generan contenidos basados en textos breves, de un máximo de 280 caracteres —hasta noviembre de 2018 eran 140—, sobre cualquier tema y en tiempo real.

De hecho, ha cobrado tal relevancia en la comunidad científica que, su equipo ha puesto a disposición una API (Interfaz de Programación de Aplicaciones) que permite extraer datos de interés. Eso sí, quizás el mayor problema al que se enfrenta Twitter son los trolls y los bots. De ahí que sea necesario mencionar el tema de la ciberseguridad, pues las organizaciones más entendidas en el ámbito del data-driven utilizan marcos y políticas para fomentar y regular el comportamiento de los usuarios digitales. De cualquier forma, para atenuar este problema, se utiliza la minería de datos

Entonces, con este entendimiento, la combinación de Twitter y la cultura data-driven podría suponer una oportunidad para las distintas organizaciones de la economía, sobre todo para el sector público, que es el caso en cuestión.

Y es que indudablemente, esta transformación digital ha sido empujada en gran parte por la Covid-19, que, dicho sea de paso, no solo desencadenó la peor crisis de la época contemporánea, también es responsable de la inmersión progresiva en el teletrabajo. Esto, sumado al confinamiento y a la necesidad del ser humano de hacerse notar a través de las redes sociales, explica porqué este es el momento apropiado.

En las sociedades emergentes, los gobiernos locales y sus respectivas empresas públicas enfrentan desafíos comunes para superar la pobreza y la exclusión social. Por esa razón, formulan y ejecutan políticas públicas orientadas a la erradicación de estos problemas. Como es bien sabido, la forma en que se abordan estos asuntos se ha volcado de lleno hacia las redes sociales. Es más, el debate público se ha instaurado principalmente en Twitter. En el caso de Ecuador, los usuarios más habituales se encuentran en las ciudades más importantes, tales como Quito y Guayaquil. Por tanto, cabe pensar que es el escenario propicio para poner a prueba la percepción de la ciudadanía y demás actores involucrados.

Volviendo a lo anterior, dado que las ciudades mencionadas son referentes en cuanto a iniciativas de desarrollo, es menester que se estudie el modo en que se construyen comunidades digitales desde las cuentas de Twitter correspondientes a sus gobiernos. Adicionalmente a ello hay que tener en cuenta las temáticas tratadas en sus agendas públicas, puesto que no basta con solo generar contenido, aunque hacerlo tenga sus ventajas.

Por otro lado, considerando que Quito y Guayaquil son ciudades grandes que plantean convertirse en inteligentes y sostenibles, podría hacer pensar que las ciudades más pequeñas no tienen esa oportunidad. Mas está claro que el tamaño de la ciudad importa, pero no es una condición esencial para volverse una smart city, según (IESE Business School, 2019). Así pues, si se aprovecha las bondades de la tecnología, cualquier otra ciudad del Ecuador que favorezca un entorno digital podrá ser objeto de investigaciones futuras e incluso entrar en algún índice internacional. Especialmente aquellas que cuenten con una buena gobernabilidad y población educada.

Esta experiencia de investigación surgió como respuesta a los problemas estructurales que acarrea el país que, por cierto, fueron más que evidentes con el inicio de la pandemia de Covid-19.

En medio del confinamiento, fue perceptible que en Ecuador se construye una sociedad funcional al establishment. Cabe recordar que, quizás nunca antes se presentó una gran oportunidad para acelerar la transformación digital. Lo grave es que durante la cuarentena no hubo medidas efectivas para mejorar el background de la población, aun sabiendo que una buena parte de ella representa la fuerza laboral nacional.

Hay un dato proporcionado por (Hootsuite & We Are Social, 2020) al cual se le pudo sacar el máximo provecho: alrededor del 70% de los ecuatorianos tiene acceso a internet y el mismo número para los usuarios de redes sociales. Los detalles sobre cómo hacerlo se discuten más adelante en este trabajo.

Aunque la pandemia agudizó las tensiones ya existentes en cuanto al manejo de la economía, con suerte, hubo quienes no perdieron su empleo y se les permitió adentrarse en el teletrabajo. Otros se decantaron oportunamente por las plataformas de e-learning y el resto, muy seguramente, se encontraba en las redes sociales.

Si bien es cierto, el gobierno avanza en asuntos de open data, pero hasta la fecha la mayoría de los GADs Municipales no incluyen esta importante temática en su agenda pública, y eso que en la actualidad, las organizaciones basan su accionar en el uso de los datos, que se conoce como data-driven culture.

Esto puede consolidarse siempre y cuando los GADs con sus empresas públicas pongan a disposición ciudadana todos sus datos relevantes, y si promueven que sus residentes reporten sobre el buen o mal funcionamiento de los servicios proveídos, participen en las iniciativas de mejora, divulguen sus opiniones sobre las decisiones tomadas, entre otros. En la coyuntura actual, esto puede representar un ahorro en gasto público de varios millones de dólares. La idea es que los gobiernos locales promuevan espacios para la interacción donde el servidor público se despoje parcialmente del conocimiento,

de modo que reinvente su relación con el ciudadano, actualmente ciudadano digital, transitando de una relación vertical a una de colaboración horizontal.

Paralelo a esto, se produjo una serie de interrogantes que buscan clarificar el escenario de Quito y Guayaquil en materia de gobernanza digital: ¿Qué tipo de relaciones se construyen desde estos GADs Municipales? ¿Cuáles son los temas prioritarios en su agenda pública? ¿Cuál es la percepción ciudadana sobre su administración?

2. Metodología

Materiales

Primeramente, para tener accesos a los datos desde de la API, fue necesario registrar una aplicación en Twitter. Una vez creada, se obtuvo las credenciales para actuar como desarrollador. R fue el software estadístico empleado durante toda la investigación.

Método

En este trabajo se presenta un Análisis Exploratorio de Datos (Exploratory Data Analysis, EDA) donde se examinan las cuentas de Twitter correspondientes a los GADs (Gobiernos Autónomos Descentralizados) de las 2 ciudades principales en Ecuador (Quito y Guayaquil), así como también las dos agencias referentes en responsabilidad social derivadas de ellos, la Agencia de Promoción Económica (ConQuito) y la Dirección Municipal de Acción Social y Educación (DASE). Adicionalmente se examinan conversaciones en base a palabras relacionadas con la temática en cuestión.

En vista de que el 29 de febrero de 2020 se confirmó el primer caso de Covid-19 en Ecuador y la cuarentena inició el 12 de marzo del mismo año, para la extracción de los tweets se ha tenido en cuenta desde el mes de junio. Esto debido a que fue

aproximadamente cuando ambas ciudades retomaron sus actividades de forma parcial. Este segundo momento se enfocó en las 4 cuentas mencionadas y recopiló 12800 tweets entre el 01-06-2020 y el 12-11-2020. Se utilizó la técnica de visualización de grafos, e histogramas para la distribución de los posts.

La parte final de este trabajo utiliza el análisis de sentimientos para aproximar la posición del ciudadano frente al accionar de los GADs mencionados. Los términos de búsqueda fueron: “alcaldiaqye” y “MunicipioQuito”, que además sirvieron para crear nubes (word clouds) con las palabras más asociadas a ellos. Se recolectaron 3200 tweets por término en un intervalo de tiempo no mayor a 7 días que es lo que permite la API de Twitter

En diversos apartados del artículo de (Ficamos & Yan, 2016) se detallan los pasos para el preprocesamiento que deben ser aplicados en esta técnica, también conocida como minería de opinión. En este otro trabajo, (Saini, Punhani, Bathla, & Kumar Shukla, 2019) enlistan los paquetes necesarios para su utilización dentro del software R.

Por otra parte se también utilizó información secundaria de acuerdo a lo reportado por (Von Feigenblatt, 2015); quien expresa que el uso de artículos académicos se recomienda como una forma de dilucidar la conexión directa entre el desarrollo de la teoría en el mundo académico y la práctica en el mundo empresarial.

3. Discusión

Bajo el supuesto de que la pandemia aceleró la transformación digital, la **Fig. 1.** muestra que después del cambio de semáforo sanitario, los posts han ido disminuyendo paulatinamente en el caso de Quito, mientras que en Guayaquil se mantienen estables. Sin embargo, las relaciones (**Fig. 2.**) que se construyen desde las cuentas son numerosas, especialmente con otras subentidades. También está

claro que existe una estrategia para el manejo de redes sociales, pero no se visualizan comunidades digitales que promuevan la utilización de datos. Por su parte, en la **Fig. 3.** es evidente que, aunque hay relaciones con otras dependencias del cabildo, las interacciones están marcadas por la presencia de influencers. En otras palabras, no es que haya una mala estrategia de redes sociales, sino que tampoco se observan ecosistemas digitales que fomenten la cultura del dato. La **Fig. 4.** muestra un nivel aceptable de percepción ciudadana, lo que se puede corroborar en la **Fig. 5.** Dicho de otro modo, la gestión del Municipio de Quito durante el período estudiado presenta scores positivos mayores que los negativos. Pero esto no significa que tenga una aprobación mayoritaria por parte de la población. Más bien parece que existe algo de temor, seguramente debido al aumento de los casos de Covid-19. En otro contexto, puede observarse que una buena parte de las palabras de la **Fig. 6.,** son manifestaciones de insatisfacción, lo cual tiene correspondencia con el análisis de sentimientos de la **Fig. 7.** donde los scores negativos superan a los positivos. Es decir, parece que el accionar de la Alcaldía de Guayaquil no es muy bien recibido por los usuarios involucrados. Ello puede atribuirse en parte a los desaciertos a la hora de tomar decisiones importantes, o, en su defecto, por los grandes retrasos en el proceso de reclamaciones lo que a su vez deriva en escaso conocimiento sobre los beneficios y servicios por parte de los ciudadanos. Este análisis proporciona información clave a la hora de dibujar una hoja de ruta accionable para los gobiernos locales, lo que puede dar lugar a la reducción de las percepciones negativas y acentuación de las positivas.

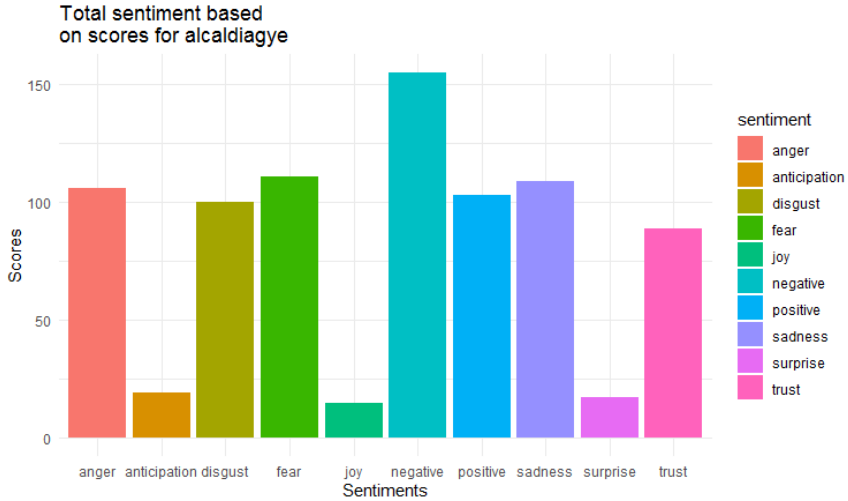


Fig. 7. Análisis de sentimientos para Alcaldía de Guayaquil

4. Conclusiones

- Un gobierno impulsado por datos no debe ser un proyecto vertical, sino que debe integrarse en la agenda de todas las instituciones gubernamentales. Lo cual no consiste en solo invertir en iniciativas de datos, sino también en la gobernanza de esos datos. Para ello se requiere la generación de una cultura del dato, donde los ciudadanos y las empresas se transformen en data citizens y data-driven companies, respectivamente.
- Social Media, en particular Twitter, es una arena ideal para los hacedores de políticas públicas. Esto dado que permite monitorear comportamientos y opiniones de los ciudadanos antes, durante o después de la implementación de una política pública. Más aún, teniendo en cuenta que este medio

provee información económica inmediata, y los perfiles son públicos por defecto.

- El análisis de Social Media puede complementar la analítica tradicional, pero no significa que vaya a reemplazar otros métodos de investigación. Como parte de futuras investigaciones, se puede incluir el análisis basado en hot spots o puntos calientes, lo cual proporcionaría información sobre los sentimientos específicos de las localidades.
- En el momento en que la ciberseguridad se afiance en Ecuador, los gobiernos locales podrían idear un plan de recompensas para que los ciudadanos participen activamente –no solo en Twitter- y reduzcan su carga tributaria (aunque podrían hacerse con otros beneficios).

5. Referencias

- Acker, A., & Clement, T. (2019). Data Cultures, Culture as Data – Special Issue of Cultural Analytics. *Journal of Cultural Analytics*, 3.
- Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a data-driven organisation. *Research-in-Progress Papers*, (pág. 43).
- BID. (2020). *Banco Interamericano de Desarrollo* . Obtenido de <https://publications.iadb.org/es/publications/spanish/document/Reporte-Ciberseguridad-2020-riesgos-avances-y-el-camino-a-seguir-en-America-Latina-y-el-Caribe.pdf>
- Ficamos, P., & Yan, L. (2016). A Topic based Approach for Sentiment Analysis on Twitter Data. *International Journal of Advanced Computer Science and Applications*, 201-204.

- Hootsuite & We Are Social. (30 de Enero de 2020). *DataReportal*.
Obtenido de <https://datareportal.com/reports/digital-2020-global-digital-overview>
- IESE Business School. (10 de Mayo de 2019). *IESE Blog Network*.
doi:DOI: <https://dx.doi.org/10.15581/018.ST-509>
- Kremser, W., & Brunauer, R. (2019). Do we have a Data Culture?
En P. Haber, T. J. Lampoltshammer, & M. Mayr, *Data Science – Analytics and Applications* (pág. 84).
- Saini, S., Punhani, R., Bathla, R., & Kumar Shukla, V. (2019).
Sentiment Analysis on Twitter Data using R. *2019 International Conference on Automation, Computational and Technology Management (ICACTM)*, (págs. 69-70).
- Sánchez Holgado, P., Martín Merino, M., & Blanco Herrero, D. (2020). Del data-driven al data-feeling: análisis de sentimiento en tiempo real de mensajes en español sobre divulgación científica usando técnicas de aprendizaje automático. *Disertaciones*, 39.
- Von Feigenblatt, Otto F. (2015). INTEGRATING THEORY AND PRACTICE IN BUSINESS EDUCATION. *Revista de Comunicación de la SEECI*, (36),183-187. Disponible en: <https://www.redalyc.org/articulo.oa?id=5235/523552855010>