

Are you watching closely? politicxsentwritteR as a magnifying glass

Gonzalez Guadalupe Andrea^a

^a Universidad de Buenos Aires

Uno de los mayores inconvenientes cuando nos proponemos realizar análisis con datos, es la recolección y la limpieza de los mismos. En este sentido, “politicxsentwritteR” es un ecosistema de herramientas realizadas en R que se propone facilitar este proceso para la producción de conocimiento -de lo que sucede en el universo twittero- sin barreras de acceso de ningún tipo.

PoliticxsentwritteR consta de una Shiny app que permite la descarga abierta de data reutilizable de los tweets emitidos por más de 500 figuras públicas de alto rango, y también de personajes de interés que coquetean constantemente con la política nacional Argentina. Además de tener una web, se creó un paquete en R donde se tiene acceso a la misma información sin necesidad de conectarse a la app.

Shiny App 1 | R package | Twitter 3 | Ciencia Política

Uno de los mayores inconvenientes cuando trabajamos con datos de twitter es el tiempo. Una vez que pasó un periodo de tiempo determinado dado un evento de interés, los usuarios tienen que pagar para obtener la data mediante la API. Esto tiene todo el sentido de ser al hablar de una plataforma privada que determina los usos y condiciones; pero ¿qué pasa cuando hablamos de data que involucra personajes políticos públicos que toman o afectan a la toma de decisiones en el ámbito público, y que a veces las comunican a través de Twitter?

Simplemente no parece adecuado perder esa data. Entonces, ¿qué pasa si derribamos esas barreras sólo para algunas figuras públicas de relevancia? Con esa idea en mente surge “politicxsentwritteR”. Un ecosistema en R que busca ayudar a la descarga de datos para más de 500 funcionarios públicos y que tiene la ventaja de que la pueden usar tanto las personas que saben R (a través del paquete) como aquellas que no (a partir de la app).

Asimismo, no se trata sólo de la descarga de datos en los tiempos permitidos por la API, sino de facilitar el análisis exploratorio a investigadores, periodistas y entusiastas en Argentina. Ambas herramientas ayudan a la ejecución de análisis exploratorios rápidos sobre la performance en Twitter de cada una de estas figuras con la intención de que buena parte del trabajo no esté enfocado en elaborar estadísticas descriptivas básicas.

Herramientas y análisis

Tanto la App como el paquete se alimentan de bases de datos que surgen de un proceso en R que está automatizado de manera diaria y que sube la información a un repositorio open source, en este caso, mongoDB como muestra la figura 1. Luego, la app y el paquete se alimentan de esas bases.

Antes de arrancar vale la pena la distinción entre figuras públicas y usuarios de twitter. Los primeros son personas físicas y los segundos son las cuentas oficiales que las utilizan para comunicar y expresar opiniones y que no siempre están

controladas ni redactadas por estas personas físicas. En este sentido, la descarga y el análisis que uno puede inferir usando la app o el paquete es acerca de los que escriben estos usuarios y no sobre las personas físicas per se.

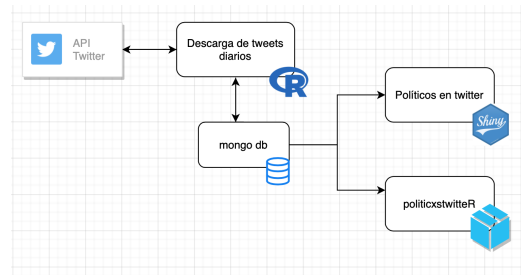


Fig. 1. Workflow de la app y el paquete en interacción con MongoDB.

La información se puede descargar por usuario o por categoría de usuario:

a. usuario: corresponde al username / nombre del account del usuario

b. categoría de usuario: existen 6 categorías correspondientes al lugar que ocupan para el mes de mayo 2021 en órganos ejecutivos, legislativos cada una de estas figuras, entre otros:

1. Todos: se refiere a todos los usuarios de los cuales se tiene información en la app.

2. Diputados: aquellos que pertenecen a la Cámara Baja del Congreso de la Nación Argentina.

Información Importante

Resulta imprescindible pensar a Twitter como un ámbito donde se discuten múltiples cuestiones de manera pública, entre ellas se discuten cuestiones de la esfera pública en términos de Habermas (1974). Y en ese sentido, es un terreno que merece nuestra atención en la medida en que nuestros representantes argumentan y debaten sobre problemáticas que afectan a la sociedad en general. PoliticxsentwritteR nace como esta posibilidad de recolectar información y hacer reusable data -que eventualmente deja de ser gratuita o accesible mediante la plataforma Twitter- de aquellos que toman o impactan sobre las decisiones en el ámbito público. Además de ser una base de datos, politicxsentwritteR nace con el objetivo de que se puedan hacer análisis en profundidad siendo que mucho del análisis exploratorio previo ya está en parte realizado mediante la Shiny app. De esta manera, investigadores pueden saltarse esa parte de ese primer análisis para trabajar en la producción de conocimiento. .

² Por cualquier duda escribir a guadagonz12@gmail.com

3. Senadores: aquellos usuarios que pertenecen a la Cámara Alta del Congreso de la Nación Argentina.

4. Poder Ejecutivo Nacional: aquellas cuentas que ocupen un lugar en el Poder Ejecutivo Nacional ya sea como presidente, vicepresidente, ministros, jefe de gabinetes y secretarios de alto rango.

5. Provincial: aquellos usuarios que se consideran importantes a tener en consideración no solamente porque ocupan cargos en la gobernación o vicegobernación sino al interior de los gabinetes de algunas provincias (CABA y Buenos Aires) y aquellos usuarios que se consideran trascendentes en la discusión pública nacional aún ocupando cargos en legislaturas provinciales (tal es el caso de Amalia Granata y Daniel Lipovetsky).

6. Otros: usuarios que forman parte del universo político a nivel nacional pero no tienen un cargo formal dentro del Estado. Muchos forman parte de la oposición, son ex funcionarios e incluso personajes que coquetean con la posibilidad de ocupar cargos públicos.

A. App. Esta app fue realizada en el marco del Grupo de Investigación en Ciencia Política “Observatorio de Redes” con el apoyo de Democracia en Red y el National Democratic Institute con el objetivo de generar una herramienta accesible para cualquier persona que quiera trabajar con datos de políticos en twitter pero que no tenga el conocimiento técnico para realizarlo. En este sentido, nació como un repositorio y puede ser descrita a partir de 3 de sus solapas:

1. La descarga de los datos: existe la posibilidad de descargar toda la data necesaria para realizar análisis sobre el comportamiento de figuras públicas, ya sea en csv como en excel.

2. Análisis de diferentes variables en un dashboard: se puede chequear la performance de los usuarios en un periodo de tiempo determinado a partir de ver el progreso en la cantidad de favoritos y retweets, en la cantidad de followers, en las nubes de palabras y en la frecuencia de los emojis que más utiliza.

3. Utiliza las redes para conocer cuánto se retweetean entre sí cada sector partidario.

Pueden acceder a la misma en este link: <https://oderedes.shinyapps.io/politicxsentwitter/>

B. R package: politicxsentwitterR. Además de una herramienta apta para todo público como es la shiny, se buscó generar una manera de poder trabajar esos datos directamente en R sin necesidad de pasar por la app y así hacerlo accesible para toda la comunidad de usuarios de R. De esta manera nació el paquete de “politicxsentwitterR” que permite realizar las mismas operaciones que la app, pero desde R y que consta de 3 funciones principales:

1.”*get_timeline_data()*” qué nos devuelve los tweets emitidos por el funcionario público para un periodo de tiempo seleccionado o el histórico.

2. ”*get_friends_followers()*” que nos descarga una evolución diaria de la cantidad de friends, followers y estados para un usuario o grupos de usuarios

3.”*get_network_data()*” que devuelve la cantidad de retweets que diferentes funcionarios se hicieron entre sí para un periodo de tiempo indicado. Repositorio del paquete: <https://github.com/guadag12/politicxsentwitterR>

Conclusiones

Con el objetivo de tener un repositorio con datos abiertos y en real time sobre lo que twittean diferentes figuras públicas relacionadas con la política nacional se creó políticos en twitter, que consta de una app y un paquete en R. La primera está pensada para cualquier persona que desee descargar los datos de tweets emitidos por diferentes usuarios, la cantidad de followers y friends así como las redes de quienes se retweetean entre sí en su computadora. La segunda para aquellas personas que programen en R y quieran obtener esos mismos datos de una manera más directa.

Agradecimientos. Nada de todo este proyecto hubiera sido posible sin Juani Belbis que me tiro la idea inicial en algo que se convirtió en uno de los mejores y más gratificantes proyectos en los que he estado involucrada hasta día de hoy. Y que, además, me inspiró para generar la shiny y colaboró y escuchó cada idea extraña que iba surgiendo en el camino; dándole soporte institucional con el Observatorio de Redes. Asimismo, resultó vital el acompañamiento de Democracia en Red (con Agustin Frizzera y Luciana Vega al timón) para el proyecto y al National Democratic Institute por interesarse en la iniciativa (de la mano de Andrea Fernandez). No quería dejar de agradecer también a mis grandes amigos Karina Bartolomé y Rafael Zambrano por ayudarme a debuggear algún que otro inconveniente surgido en el camino y cuyo apoyo e impulso hizo que este proyecto se convirtiera en algo más grande de lo que había pensado en primer lugar.

References

Habermas, J. (1974). The public sphere: An encyclopedia article. (1964). New German Critique, 3, pp. 49-55.