Dirección de Desarrollo Digital

Unidad de Científicos de Datos

El futuro es de todos

DNP
Departamento
Nacional de Planeación

ANÁLISIS DESCRIPTIVO DE TWEETS DE LOS EQUIPOS DE RESPUESTA ANTE EMERGENCIAS INFORMÁTICAS

Entidad
Departamento Nacional de Planeación

- Dirección de Desarrollo Digital.
- Oficina de Tecnologías y Sistemas de Información

Sector	Lenguaje	Fuente de datos
TIC	Python	Twitter

Presentación
Los incidentes de seguridad de la información son eventos fortuitos que ponen en peligro los datos cibernéticos. Por fortuna, Los Equipos de Respuesta ante Emergencias Informáticas (o CSIRT por sus siglas en inglés) se encargan de tomar acciones preventivas y correctivas ante los riesgos que se presentan a nivel mundial, siendo Twitter uno de los canales de información que estos equipos utilizan para transmitir las alertas. Para la Oficina de Tecnologías y Sistemas de Información (OTSI) es importante realizar un seguimiento a estos riesgos, ya que identificar las vulnerabilidades oportunamente les permitirá tomar mejores decisiones en cuanto a la seguridad de la información en el DNP. Dado lo anterior, la OTSI acudió a la Unidad de Científicos de Datos (UCD) con el propósito de implementar una herramienta que facilite la consulta de tweets y permita reducir el tiempo de búsqueda, análisis e identificación de términos clave en los tweets.

Information security incidents are events that put cyber data at risk. Fortunately, Computer Security Incident Response Teams (or CSIRTs) are responsible for taking preventive and corrective actions against risks that occur worldwide, with Twitter being one of the information channels that these teams use to transmit alerts. For the Office of Information Technologies and Systems (OTSI) it is important to monitor these risks, as identifying vulnerabilities in a timely manner will allow them to make better decisions regarding information security in the DNP. Given the above, OTSI turned to the Data Scientists Unit (UCD) with the purpose of implementing a tool that facilitates the consultation of tweets and allows for a reduction in the time required to search, analyze and identify key terms in the tweets.

Objetivo general
Implementar una herramienta de consulta que permita identificar las palabras claves más usadas en los tweets de diversos equipos de respuesta a incidentes (CSIRT).

Objetivos específicos

1. Crear un script que permita descargar y actualizar los tweets de los usuarios de interés.
2. Implementar un script para la lectura, limpieza y procesamiento de los Tweets.
3. Desarrollar una aplicación web que permita consultar los resultados del análisis.

Metodología

La metodología planteada para la ejecución del proyecto puede resumirse en cuatro etapas: (1) Descarga de la información de Twitter; (2) Procesamiento de texto de los tweets; (3) Identificación de términos clave y (4) Visualización de los resultados.

Descarga de la información de Twitter

Con el fin de obtener los tweets de los usuarios CSIRT de interés, se usó la librería *Tweepy*, la cual permite conectarse a la API (Interfaz de Programación de Aplicaciones) de Twitter, para esto se requiere contar con unos tokens asociados a una cuenta de desarrollador de Twitter. Para la descarga se utilizó un listado de usuarios de interés previamente seleccionados por la OTSI, el script recorre la lista de usuarios de forma iterativa y permite descargar el texto de los tweets junto con atributos adicionales como son el número de *likes*, número de *retweets*, *hashtags*, *links* de acceso, fecha y hora de publicación, entre otros. Una vez se tiene la información, esta es exportada a un archivo *XLSX,* el cual servirá de insumo en el procesamiento y visualización de resultados. Vale la pena mencionar que la base de datos solo almacena los tweets de los últimos 90 días ya que el enfoque del proyecto no considera mantener una base de datos extensa de estos.

Procesamiento de texto de los tweets

A partir de la información descargada, se procedió a realizar una limpieza al texto de los tweets, en primera medida se removieron elementos ajenos al texto en sí, como son hashtags, menciones de usuarios y links externos. Esto se hace para remover ruido en el análisis de texto a realizar en las etapas siguientes.

En segundo lugar, se detectó el lenguaje y se eliminaron los *stopwords* de acuerdo con el idioma del tweet. Los *stopwords* son palabras que no aportan valor al análisis descriptivo, estas palabras son usualmente preposiciones, conectores y nombres propios como “de”, “a”, “la”, entre otros. Por último, se transformó el texto a minúsculas y se retiraron signos de puntuación, números, espacios dobles y tildes con el fin de estandarizar las palabras y facilitar la comparación entre estas, es decir, para que palabras como “Informática” e “INFORMATICA” no se consideren distintas, sino que se utilice únicamente la palabra “informatica”. El procedimiento de limpieza descrito anteriormente se realizó con la librería ConTexto, en la Tabla 1 se presentan dos ejemplos del procesamiento realizado al texto.

Tabla 1: Procesamiento de tweets.

Tweet	Tweet sin menciones	Tweet final
¿Tienes dudas sobre #ciberseguridad y no sabes a quién acudir? Desde @INCIBE queremos recordarte que estamos a tu disposición en nuestro teléfono gratuito ☎️7. ¡Te ayudaremos! https://www.osi.es/es/contacto pic.twitter.com/uJ3zNQxfda"	¿Tienes dudas sobre y no sabes a quién acudir? Desde queremos recordarte que estamos a tu disposición en nuestro teléfono gratuito ☎️7. ¡Te ayudaremos! 	dudas sabes quien acudir queremos recordarte disposicion telefono gratuito ayudaremos
Recuerda que en nuestra web encontrarás varias campañas de concienciación con las que descubrir buenas prácticas en #ciberseguridad. ¡No lo pienses más! #OSIconcienciación https://youtu.be/Djw0MNmdB7Y	Recuerda que en nuestra web encontrarás varias campañas de concienciación con las que descubrir buenas prácticas en ¡No lo pienses más!	recuerda encontraras varias campanas concienciacion descubrir buenas practicas pienses

Fuente: Elaboración propia.

Se debe mencionar que el proceso de limpieza busca estandarizar las palabras para facilitar su comparación y análisis, sin embargo, también puede introducir errores particularmente en el idioma español al remover signos de puntuación

y caracteres especiales, en la Tabla 1 se puede observar como se modifica la palabra "*campañas*" obteniendo como resultado la palabra "*campana*", lo cual podría llegar a ocasionar errores en la interpretación de resultados.

Identificación de términos clave

Luego de obtener los tweets procesados y los hashtags, se identificaron las palabras que presentaron una mayor frecuencia en los textos, así como también los hashtags más utilizados. Para esto, se procedió a generar n-gramas o conjuntos de n palabras seguidas donde n es un número entero mayor a cero, a partir de los textos de los tweets y contar la frecuencia de aparición de estos. A manera de ejemplo, al generar los unigramas (n-gramas donde n es igual a 1) del texto "*Identificación de riesgos informáticos*", obtendríamos 4 unigramas correspondientes a cada palabra, es decir, "Identificación", "de", "riesgos", "informáticos", y al generar los bigramas (n-gramas donde n es igual a 2) se obtienen 3 términos, "Identificación de", "de riesgos", "riesgos informáticos".

En el caso de los hashtags, solo se calcula su frecuencia en los diferentes tweets, mientras que para el resto del cuerpo del tweet se distingue la frecuencia de unigramas y bigramas (secuencia de dos palabras), una vez identificados los términos más frecuentes, se procedió a implementar la visualización de estos mediante gráficos de barras de frecuencias y nubes de palabras.

Visualización de los resultados

Para la consulta de tweets y visualización de los términos identificados se utilizó el framework *Streamlit,* este permite implementar de manera sencilla una interfaz web de consulta de información basada en *Python.*

La aplicación contiene varios elementos, siendo el primero una tabla que muestra los tweets junto con sus atributos de fecha y hora de publicación, número de *likes*, número de *retweets*, usuario que hizo la publicación y un enlace para consultar el *tweet* en Twitter, facilitando la consulta todos los tweets de usuarios de interés de manera consolidada. Esta tabla se puede organizar de acuerdo con los atributos mencionados, lo que facilita consultar los *tweets* más recientes, antiguos, o los que presentan un mayor número de *likes* o *retweets*. En la Figura 1 se presenta una imagen de la tabla de *tweets* implementada.

Figura 1. Tabla de Tweets

Tabla de Tweets

Ordenar tweets por:

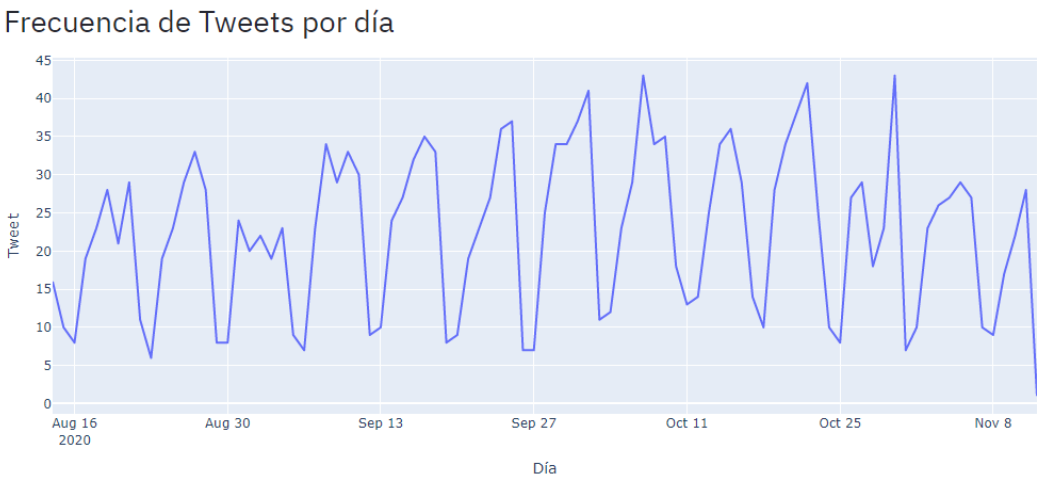
Recientes

Fecha	Usuario	Tweet	Visualizar	Likes	Retweets
12 Nov 2020 01:25 AM	bjCSIRT	Deuxième nuit d'effort pour le sacre final dans 10h. #HackerLab2020 https://t.co/dSW63uLLdJ	Ver Tweet	3	2
11 Nov 2020 08:34 PM	osiseguridad	Nuestra nueva guía de #ciberataques te trae todo lo que debes saber a nivel usuario sobre los #virus y #malware. Descárgala y conoce las principales amenazas. 📥Disponible en: https://t.co/nlxlvS8Nya https://t.co/wkMViDcMnX	Ver Tweet	9	9
11 Nov 2020 07:17 PM	osiseguridad	¿Sabrías distinguir un #hacker 🕵 de un #ciberdelincuente 🕵? ¿Cuáles son sus características y las motivaciones para actuar? Sal de dudas con este interesante vídeo de @osiseguridad https://t.co/GONmF1uvQ1	Ver Tweet	89	38
11 Nov 2020 06:35 PM	CSIRT_Telconet	Fallos severos en Cisco causarían denegación de servicio en varios de sus productos de seguridad https://t.co/wpl9sCjyXS https://t.co/19x64rNVem	Ver Tweet	0	0
11 Nov 2020 06:07 PM	bjCSIRT	#Bénin : c'est parti pour l'acte 3 du #HackerLab de l' @Anssi_Benin Comprendre le concept🔻 https://t.co/bQwhcZwRky @numeriquebenin @adnbenin @arcepbenin @gouvbenin https://t.co/NkSqLCcmzi	Ver Tweet	11	4
11 Nov 2020 06:07 PM	bjCSIRT	Mens sana in corpore sano !! Challenge spécial :) !! https://t.co/XR40HhWJ5H	Ver Tweet	18	3
11 Nov 2020 06:06 PM	osiseguridad	El #spoofing es una técnica de #suplantación de identidad muy común. ¿Sabrías cómo identificar un "mail spoofing"? ➡Descubre cómo: https://t.co/NTU5sXBWub https://t.co/ZJ0xhuB7sD	Ver Tweet	8	8
11 Nov 2020 06:06 PM	bjCSIRT	Alors que la dernière nuit de hacking s'annonce, les ardeurs sont éprouvées mais les challengers se donnent le courage avec une série de pompes https://t.co/8eDjidvjXC	Ver Tweet	14	5

Fuente: Elaboración propia.

Seguida de esta, en la Figura 2 se presenta un gráfico de líneas con la cantidad de Tweets publicados por día, por medio de esta gráfica se espera que la OTSI pueda identificar un riesgo o evento que tenga gran impacto, teniendo en cuenta que se espera que aumente significativamente el número de tweets, por ejemplo, en el 2018 con la aparición del virus tipo ransomware WannaCry. Se evidencia un patrón en la cantidad de tweets publicados en cuanto el número de estos sube y baja, las "bajadas" o la reducción del número de publicaciones corresponden a los sábados y domingos.

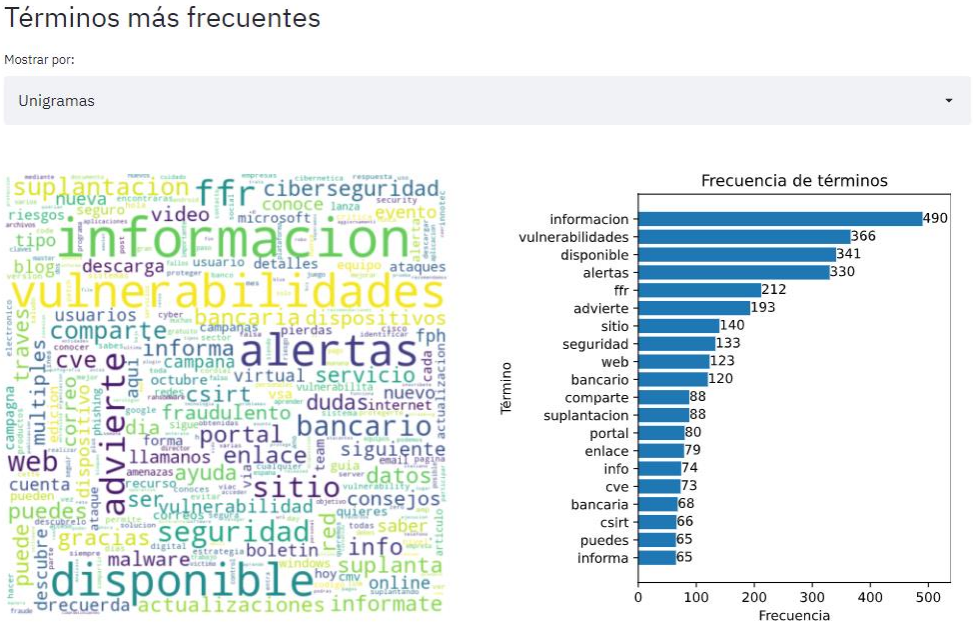
Figura 2. Número de Tweets publicados por día.



Fuente: Elaboración propia.

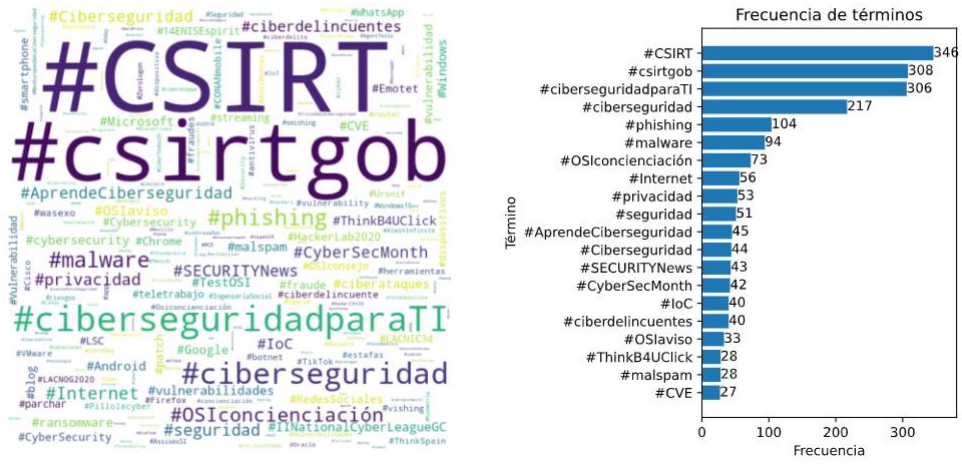
En cuanto a los términos más frecuentes se hace una representación por unigramas o bigramas según seleccione el usuario, utilizando una nube de palabras en la que se presentan hasta 300 términos y en la que el tamaño de cada palabra se asocia al número de veces presente en los textos, es decir, a mayor frecuencia de ocurrencia mayor será el tamaño de la palabra. Al lado de esta se presenta un gráfico de barras con los 20 términos más frecuentes en los textos. En la Figura 3 y Figura 4 se presentan los gráficos de términos y hashtags más frecuentes respectivamente. Estas gráficas permitirán comprender en mejor medida y a nivel general el contenido de los textos de los tweets.

Figura 3. Términos más frecuentes.



Fuente: Elaboración propia.

Figura 4. Hashtags más frecuentes.



Fuente: Elaboración propia.

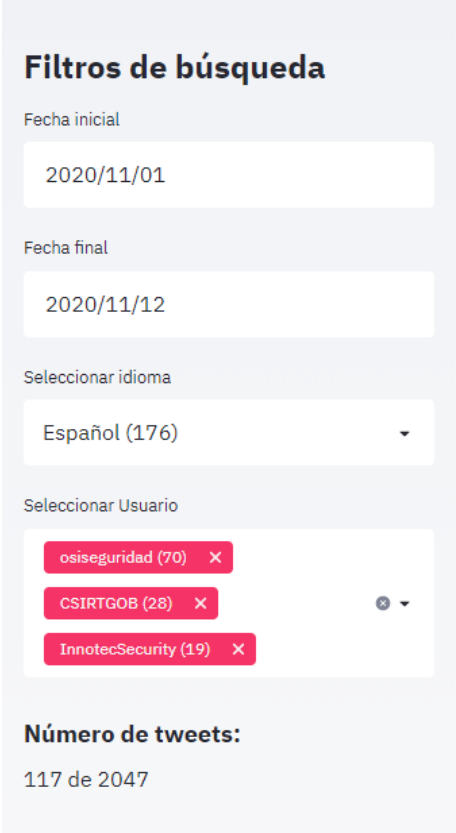
Una vez se inició el desarrollo del proyecto se pudo identificar que muchos de los tweets contenían imágenes, por lo que un análisis de texto podría quedar corto, teniendo en cuenta lo anterior se agregó un carrusel de imágenes para facilitar la consulta de estas. En la sección de imágenes se presenta en paréntesis junto al título "Galería de imágenes" el número de imágenes disponibles para consultar, la imagen contenida en el *tweet*, el texto del *tweet* y un enlace de acceso con el texto "Ver Tweet" que permite consultar el tweet en la página de Twitter, de este modo se puede consultar rápidamente el conjunto de imágenes adjuntas a los *tweets* descargados.

Figura 5: Galería de imágenes

Fuente: Elaboración propia.

Adicional a los componentes mencionados, se tiene una barra al margen izquierdo la cual permite filtrar el conjunto de *tweets* a analizar, teniendo en cuenta un rango de fechas según la fecha de publicación, el usuario que hace la publicación y el idioma predominante en el *tweet*. En la Figura 6 se presenta un ejemplo de los filtros de búsqueda, en este caso se hace un filtro por fechas, limitando la consulta entre el 01 y 12 de noviembre del año 2020, de los tweets publicados en esos días se seleccionaron los que están en español, se observa un número entre paréntesis al lado de "*Español*" indicando que hay 176 tweets en este idioma, por último se seleccionan los tweets de los usuarios *osiseguridad*, *CSIRTGOB* e *InnotecSecurity*, los cuales publicaron 70, 28, y 19 tweets respectivamente. Al final de los filtros se indica el número de tweets usados en el análisis con relación al número total de tweets disponibles en la base de datos.

Figura 6. Filtros disponibles del conjunto de tweets



Fuente: Elaboración propia.

El despliegue de la herramienta de consulta se realizó en el servidor de la UCD y se puede acceder a través del enlace http://vdatascience:8065/. Vale la pena mencionar que solo se podrá acceder a esta desde la intranet del DNP.

El proceso de actualización de información se realiza mediante la ejecución programada de un script los domingos cada 7 días, el cual se encarga de ejecutar las rutinas de descarga de *tweets*, procesamiento de textos y actualización del tablero de visualización.

Resultados

La aplicación desarrollada permite visualizar los *tweets* de los usuarios de interés y mediante gráficos descriptivos entender en mejor medida el contenido de estos, permitiendo reducir los tiempos de consulta por parte de la OTSI. Con esta implementación se espera que desde la OTSI se pueda realizar un seguimiento más riguroso a los riesgos identificados por los CSIRT y así poder tomar acciones preventivas y correctivas oportunamente frente a estos.

Conclusiones y recomendaciones

1. La herramienta desarrollada facilita el análisis de tweets de los diferentes usuarios de interés, lo que reduce el tiempo de búsqueda por parte de la OTSI.
2. La herramienta sirve como un apoyo en la toma de decisiones, no sustituye el análisis de los expertos temáticos que realizan seguimiento a novedades de los CSIRT.
3. La herramienta consiste en una interfaz web para la consulta de los resultados del análisis, más no una aplicación web desarrollada con un rigor técnico ya que no cuenta con mecanismos de escalabilidad, no utiliza una base de datos y no cuenta con mecanismos de seguridad.

Socialización

La herramienta fue desplegada en un servidor del DNP para que pueda ser utilizada desde la intranet, y fue presentada a la Oficina de Tecnologías y Sistemas de Información.