# Cambridge Analytica robo de datos

## Introducción

Facebook revelo que la cantidad de usuarios de su red social en la Unión Europea cuya información fue sustraída por la firma de datos Cambridge Analytica es bastante grande y esto se prestó para formar teorías conspiratorias.

Una gran parte de los usuarios de Facebook fue afectada por Cambridge Analytica, se estima que alrededor de 2.7 millones de usuarios de Facebook en la Unión Europea fueron sustraído por la firma de datos con sede en Londres, confirmó la red social a la Comisión Europea (CE), según informó Reuters este viernes.

La CE fue notificada a través de una carta que recibió por parte de Facebook el jueves por la noche. En ella, la red social también explica los pasos que ha dado desde entonces. Un portavoz del organismo europeo.

El pasado miércoles, Mark Zuckerberg admitió que alrededor de 87 millones de personas en todo el mundo, la mayoría en Estados Unidos, fueron afectadas por el robo de datos de la empresa británica. En un principio, la cifra reportada fue de 50 millones, aunque Chris Wylie, ex empleado de Cambridge Analytica, ya había adelantado que ese número era mucho mayor.

Estados Unidos encabeza con más de 70 millones de personas la lista publicada por Facebook de los países con mayor número de usuarios afectados. Entre esos 10 países también se encuentra Reino Unido, el cual se ubica en la cuarta posición con al menos 1.07 millones de afectados.

En México, donde se mantienen encendidas todas las alertas rojas luego de que las campañas electorales para la presidencia del país iniciaron hace una semana, la red social ha confirmado que los datos de 789.800 mexicanos fueron comprometidos por Cambridge Analytica.

## Algoritmos utilizados por Cambridge Analytica

El algoritmo en el corazón de la violación de datos de Facebook suena casi demasiado distópico como para ser real. Analiza los mensajes más aparentemente triviales y desechables -los usuarios "me gusta" reparten mientras navegan por el sitio- para recopilar información personal sensible sobre orientación sexual, raza, género, incluso inteligencia y traumas infantiles.

Unas cuantos de "Me gusta" pueden dar una fuerte predicción de por qué partido votará un usuario, revelar su género y si su pareja es probable que sea un hombre o una mujer, proporcionar pistas poderosas sobre si sus padres permanecieron juntos durante su niñez y predecir su vulnerabilidad al abuso de sustancias. Y puede hacer todo esto sin necesidad de profundizar en mensajes personales, publicaciones, actualizaciones de estado, fotos o toda la otra información que posee Facebook.

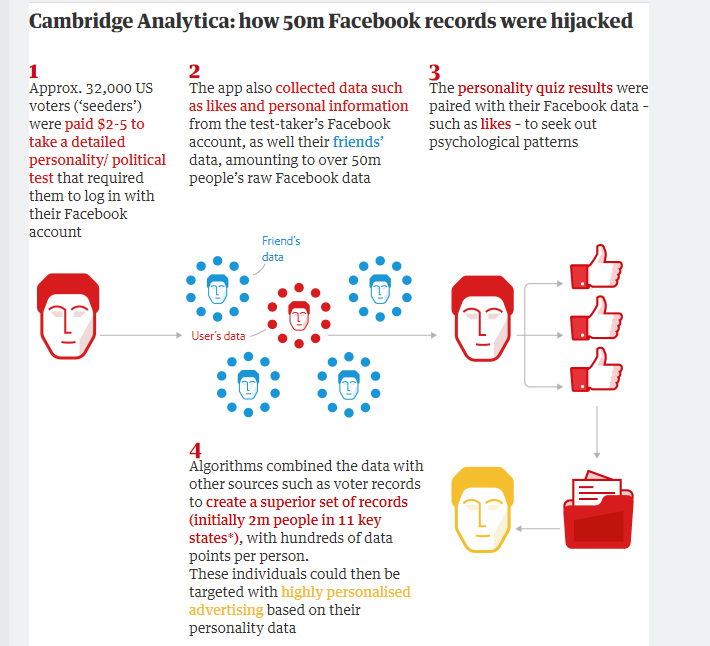
Quizás esto no suene relevante, pero estudios realizados por psicólogos mostraron que se podían deducir rasgos mucho más complejos a partir de patrones invisibles para un observador humano que escaneaba los perfiles. Solo unos pocos "me gusta" aparentemente aleatorios podrían formar la base para evaluaciones de personajes inquietantemente complejas.

### ¿Como funciona el algoritmo?

La aplicación registró los resultados de cada prueba, recopiló datos de la cuenta de Facebook del usuario y, de manera crucial, extrajo los datos de sus amigos de Facebook también.

Los resultados se combinaron con los datos de Facebook de cada encuestador para buscar patrones y construir un algoritmo para predecir los resultados para otros usuarios de Facebook. Los perfiles de sus amigos proporcionaron un terreno de prueba para la fórmula y, lo que es más importante, un recurso que haría que el algoritmo fuera políticamente valioso.

Para ser elegible para tomar la prueba, el usuario tenía que tener una cuenta de Facebook y ser un votante de EE. UU., Por lo que decenas de millones de perfiles podrían coincidir con las listas electorales. A partir de una prueba inicial de 1,000 "sembradoras", los investigadores obtuvieron 160,000 perfiles, o alrededor de 160 por persona. Eventualmente, unos cientos de miles de personas que tomaron el examen serían la clave de los datos de una amplia franja de votantes estadounidenses.



En esta imagen se muestra de una manera grafica como funciona el algoritmo utilizado por Cambridge Analityca, para la recopilación de la información de los usuarios.

## Consecuencias

En cuestión de meses, Kogan y Cambridge Analytica tenían una base de datos de millones de votantes de EE. UU. Que tenían su propio algoritmo para escanearlos, identificando probables tendencias políticas y rasgos de personalidad. Luego, podrían decidir a quién orientar y elaborar sus mensajes que probablemente les atraiga a ellos, un enfoque político conocido como "micro-targeting".

Facebook anunció el viernes que suspendía Cambridge Analytica y Kogan de la plataforma en espera de información sobre el uso indebido de los datos relacionados con este proyecto.

## Conclusiones

Facebook negara que la recolección de decenas de millones de perfiles por parte de GSR y Cambridge Analytica hayan sido una violación de datos.

Dijo en un comunicado que Kogan "obtuvo acceso a esta información de manera legítima y a través de los canales adecuados", pero "no acató nuestras reglas" porque transfirió la información a terceros.

Este evento que sucedió entre Facebook y Cambridge Analityca, no solo afecta a EEUU o a la Unión Europea, ya que En México, donde se mantienen encendidas todas las alertas rojas luego de que las campañas electorales para la presidencia del país iniciaron hace una semana, la red social ha confirmado que los datos de 789.800 mexicanos fueron comprometidos por Cambridge Analytica.

Las elecciones presidenciales del próximo 1 de julio en México podrían convertirse en un gran reto para el país. Las campañas han empezado cuando el escándalo de Cambridge Analytica está en su punto más álgido y en un momento en el que Facebook está pasando por una crisis de confianza por las políticas de protección de usuarios.

En México, las denuncias por compra de votos en período electoral o por otros intentos de interferir en el resultado de los comicios se han presentado en más de una ocasión, pero, en 2018, el país está haciendo frente a lo que podría ser una crisis de seguridad cibernética. Lejos de ponerle freno, las evidencias de la actividad de Cambridge Analytica en el país se unen a filtraciones de datos que ponen el panorama todavía más complicado.

"Como consecuencia de esas filtraciones, el INE ha tomado medidas preventivas para que, sin demérito de los derechos de los partidos políticos establecidos en la ley para revisar las bases de datos del Padrón Electoral, se impida el mal uso de los datos contenidos en el mismo", señaló Lorenzo Córdova, consejero presidente del órgano electoral.

El PRI fue multado con 43.7 millones de pesos, mientras que el Movimiento Ciudadano debe hacer frente a una sanción de 34.2 millones de pesos. La medida tomada por el INE tiene como objetivo prevenir que otros casos parecidos tengan lugar y que los datos sean utilizados de manera ilícita.

Pero ¿para qué se pudieron haber utilizado estos datos? El caso de Cambridge Analytica y el Russiagate pusieron de relieve cómo la información personal puede ser utilizada para influir en unas elecciones como las de Estados Unidos de 2016. Los dos escándalos demostraron lo fácil que puede llegar a ser utilizar los datos obtenidos para favorecer a un candidato.

Además, Luis Fernando García señaló que la utilización de la información para fines publicitarios es otro de los objetivos de estas filtraciones. Los datos que pueden ser útiles para dirigir anuncios a un público en concreto y puede ser una gran fuente de ingresos. Asimismo, obtener información personal de personas físicas puede ser la mejor manera de extorsionarlas o de suplantar sus identidades

Las campañas presidenciales ya han empezado y, con ellas, la difusión de las noticias falsas, los intentos de desprestigiar a políticos y, posiblemente, las estrategias para poder influir en el voto de los ciudadanos. Asociaciones como R3D están investigando la situación para poder identificar problemas de seguridad cibernética aunque, desde el lado gubernamental, los expertos critican la falta de medidas drásticas que podrían poner fin a la filtración y utilización de datos personales para fines electorales.