

# Propuesta de TFG

Javier Sánchez Mendoza

Grado de ingeniería informática

Health IT

Carlos Luis Sánchez Bocanegra

José Antonio Morán Moreno

7-3-2018

## TFG - Health IT: Propuesta de TFG

<b>Propuesta de TFG .....</b>	<b>2</b>
<b>Título .....</b>	<b>2</b>
<b>Palabras clave .....</b>	<b>2</b>
<b>Temática escogida.....</b>	<b>2</b>
<b>Problemática a resolver.....</b>	<b>2</b>
<b>Objetivos.....</b>	<b>3</b>
<b>Bibliografía.....</b>	<b>3</b>

## Propuesta de TFG

### Titulo

Obtención de Big Data destinado al estudio del movimiento antivacuna haciendo crawler sobre Twitter y YouTube.

### Palabras clave

Big data, crawler, Twitter, YouTube, antivacuna.

### Temática escogida

Desde la introducción de la vacunación como método preventivo de enfermedades han existido entidades y grupos de personas que se han opuesto a ella y han dudado de su efectividad o propósito [1]. Hoy en día el activismo anti-vacunación (conocido también como movimiento antivacunas) ha vuelto a la actualidad y se encuentra en auge en algunas regiones del planeta tales como Europa o Estados Unidos, cobrándose en el peor de los escenarios víctimas mortales a causa de enfermedades que se creían erradicadas y que han vuelto a surgir [2].

Para hacer posible el estudio y comprensión de las motivaciones del movimiento antivacuna se propone el desarrollo de una aplicación que permita la recolección de grandes cantidades de datos de la actividad realizada por parte de este colectivo en redes sociales (tales como Twitter y YouTube) con el fin de poder habilitar su posterior tratamiento y estudio por parte de un analista de datos (data scientist [3]).

### Problemática a resolver

Proyecto que se enmarca dentro de la problemática de la obtención, almacenamiento y procesamiento de grandes volúmenes de datos (Big Data [4]) procedentes de diversas fuentes.

Hoy en día las redes sociales han puesto al alcance de los analistas de datos una gran cantidad de datos disponibles para ser analizados. Una de las problemáticas a las que se quiere hacer frente es la obtención de información procedente de distintas fuentes de datos de forma homogenizada. Para ello se propone hacer uso de interfaces de programación de aplicaciones (abreviado como API [5] en inglés) ofrecidas públicamente por distintas redes sociales de tal forma que el proceso resulte transparente para el usuario final, en nuestro caso un analista de datos, permitiendo la extracción de este problema.

La obtención de grandes volúmenes de datos nos llevara también a la problemática que surge en su almacenamiento en bases de datos tradicionales y su posterior procesamiento. Para habilitar al usuario final el correcto acceso a la información obtenida se estudiaran las ventajas que aporta el uso de base de datos NoSQL [6] para este cometido, al ser diseñadas especialmente para manejar enormes cantidades de datos.

## Objetivos

Los objetivos que se quieren lograr con este proyecto son los siguientes:

- Investigar que funcionalidades aportan las API públicas ofrecidas por YouTube y Twitter y analizar como se pueden utilizar para la obtención de la información requerida.
- Permitir la obtención de información procedente de distintas redes sociales (YouTube y Twitter) de forma homogénea según criterios de búsqueda proporcionados por el usuario final.
- Habilitar la gestión, visualización y exportación de datos obtenidos en distintos procesos de extracción para su posterior análisis en herramientas especializadas.
- Proporcionar una interfaz de usuario usable que permita realizar las acciones requeridas por el usuario final (tales como la ejecución de procesos de extracción de datos, su gestión, exportación y visualización).
- Determinar como almacenar y acceder de forma eficiente a la gran cantidad de información que será obtenida.

En resumen, el principal objetivo del proyecto es el de proporcionar una aplicación que permita a un analista de datos obtener, de forma usable y transparente, la información que requiera de las redes sociales YouTube y Twitter enfocado a poder realizar una investigación sobre el movimiento antivacunas.

## Bibliografía

- [1] [https://es.wikipedia.org/wiki/Controversia\\_de\\_las\\_vacunas](https://es.wikipedia.org/wiki/Controversia_de_las_vacunas) (07/03/2018)
- [2] <http://www.elmundo.es/cataluna/2015/06/27/558e5fb2e2704ea41e8b4576.html> (07/03/2018)
- [3] [https://es.wikipedia.org/wiki/Ciencia\\_de\\_datos](https://es.wikipedia.org/wiki/Ciencia_de_datos) (07/03/2018)
- [4] <https://es.wikipedia.org/wiki/Macrodatos> (07/03/2018)
- [5] [https://es.wikipedia.org/wiki/Interfaz\\_de\\_programación\\_de\\_aplicaciones](https://es.wikipedia.org/wiki/Interfaz_de_programación_de_aplicaciones) (07/03/2018)
- [6] <https://es.wikipedia.org/wiki/NoSQL> (07/03/2018)