

Con el objetivo de facilitar el proceso de desarrollo y la colaboración interdisciplinaria del equipo es necesaria la creación de un sistema que permita recopilar, almacenar, consultar y analizar la información. Para ello, se utiliza la API pública de Twitter, bases de datos relacionales, diferentes niveles de abstracción y una aplicación gráfica.

## Recopilación de datos

El alcance de esta investigación está orientado al análisis de noticias falsas en la plataforma de Twitter. Es por tanto necesario comunicarse con la misma para obtener los datos necesarios. Una API (Application Programming Interface) permite a una aplicación comunicarse con otra mediante protocolos comprensibles por ambas. La API pública de Twitter permite obtener los datos disponibles sobre cada tweet y cada usuario, así como realizar búsquedas complejas. Esto sirve entonces como el punto de partida de la información a ser recopilada.

## Almacenamiento de datos

Una vez identificados los datos, para un mejor desarrollo de este estudio, es conveniente agruparlos (junto con otros datos adicionales que resulten del análisis) en una fuente central de información que pueda ser accedida por todos los integrantes del proyecto, así como otros componentes del sistema. Para ello se implementa una base de datos relacional, compuesta por diversas tablas con columnas que definen las propiedades de cada registro. Para esta etapa se implementan:

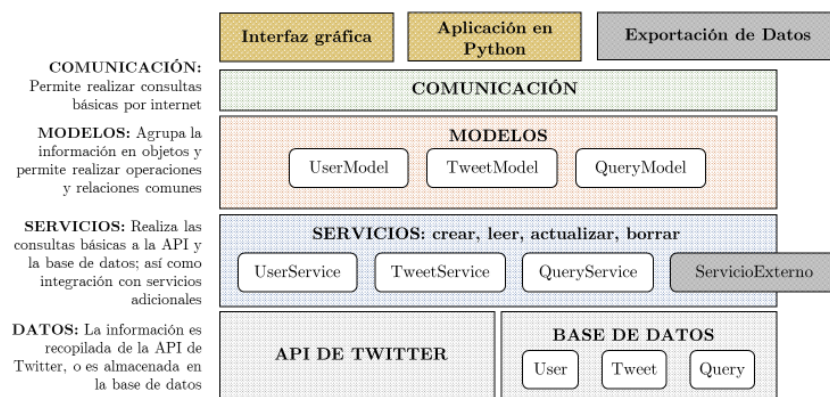
- Users: Almacena los datos de los usuarios, tales como creación, nombre y descripción de perfil. Permite la identificación de cuentas anómalas.
- UserStatsFreeze: Almacena estadísticas tales como número de seguidores, seguidos y publicaciones. Permite la identificación del alcance de una cuenta.
- UserAnalysis: Agrupa todos los datos que emanan del desarrollo de este estudio.

- Tweet: Almacena los datos de las publicaciones, así como su contenido y referencias a otras entidades (como respuestas a usuarios o tweets).
- TweetStatsFreeze: Almacena estadísticas relevantes al alcance de una publicación, como número de respuestas, favoritos y retweets.
- TweetAnalysis: Agrupa todos los datos que emanan del desarrollo de este estudio.
- Query: Permite almacenar, consultar y replicar criterios de búsqueda en la plataforma de forma periódica.
- QueryTweets: Asocia los tweets obtenidos con su respectiva consulta.
- QueryAnalysis: Agrupa los datos que emanen del desarrollo de este estudio.

## Abstracción en diferentes niveles

Debido a la amplia cantidad de interconexiones entre las entidades, así como la integración con servicios adicionales que se mencionan más adelante, las consultas pueden volverse muy complejas, así como su implementación. Separar los comandos en capas diferentes, acotadas por su responsabilidad, permite abstraer la complejidad y crear un sistema más robusto. Por ello se implementó el siguiente modelo:

Arquitectura del Sistema de Información



## Implementación de Sistema Gráfico

La anterior arquitectura permite el manejo de los datos de forma modular e integrable con el resto de los componentes del sistema. Igualmente provee una forma de construir aplicaciones de interacción de datos como nivel adicional de abstracción. Con el propósito de facilitar el análisis manual y supervisado se desarrolló una herramienta gráfica que permite visualizar de forma comprensible los datos almacenados. Esto permite también una colaboración mas estrecha entre los equipos interdisciplinarios.

El sistema presenta inicialmente la tabla Tweets descrita en los apartados anteriores. Cada una de las filas despliega un cuadro de diálogo con más información relevante. Los hipervínculos a su vez despliegan cuadros de diálogo con información relevante de los mismos. El desarrollo de este sistema de información, así como de una herramienta de visualización permite un fácil y eficiente manejo de los datos que permite el desarrollo de la solución propuesta.

The application interface consists of a main window titled "Tec Fake News Database" with a URL of <https://tec-fake-news.herokuapp.com>. The main view displays a table of tweets with columns: tweetID, authorID, inReplyToUserID, inReplyToTweetID, quotesTweetID, creationDate, and fullText. The table contains 10 rows of data.

Clicking on a tweet ID opens a "Details" dialog box. For example, clicking on tweet ID 131618723180134409 shows details for a tweet by user 76141225 (fabriziomexja) about COVID-19 vaccines. The dialog includes fields for Tweet ID, Author, Replies to, Retweets, and Full Text.

Clicking on a user ID opens a "Details" dialog box. For example, clicking on user ID 76141225 shows details for user fabriziomexja, including creation date, full name, screen name, biography, and location (Mexico). The dialog also includes a "Tweets by" section showing a tweet by the same user.