

Identificadores de usuarios, posts, tuits, etc

Ya sabemos que las redes sociales son en realidad bases de datos que organizan datos categorizados como objetos (vídeos, fotos, posts, tuits, etc). Para que estos objetos puedan ser encontrados rápidamente para mostrar al usuario cuando lo requiera, es necesario que tengan un identificador único (**id**).

Las redes sociales usan diversos tipos de ids públicos para identificar sus productos. Por ejemplo Youtube usa una serie extraña de characters como id público de sus vídeos. Este vídeo que encontré al azar tiene el identificador **846cqfUuQZA**:



Figure 1: fuente [Youtube](#)

Identificadores de Twitter

Los ids de Twitter son más simples. Son una serie de números que van aumentando uno a uno conforme se vayan creando nuevos tuits.

Por ejemplo el [primer tuit de nuestro presidente](#) tiene el id 15295318521. O sea quince mil doscientos noventa y cinco millones trescientos dieciocho mil quinientos veintiuno.

El primer tuit de la historia debería tener como id el número **uno** (o el número **cero** si eres computadora y no humano). Pero en realidad el primer tuit (reliquia histórica) pertenece a un [co-fundador de Twitter](#) y tiene el id **20**. Este es:



Figure 2: [tuit](#)

<https://twitter.com/jack/status/20>

Ventajas de usar números como ids

Cuando los programadores quieren acceder a tuits en grandes cantidades tienen que usar la interfaz diseñada por Twitter para estos propósitos (la conocida [API versión 1.1](#)). Esta interfaz es la que usan los tuiters bots y varias aplicaciones que analizan twitter con fines de marketing, investigación científica, etc.

El funcionamiento del API de twitter depende críticamente del hecho que sus ids sean números. Por ejemplo cuando un bot hace búsquedas de tuits que usan un determinado *#hashtag* puedes restringir los pedidos a twitter para que solo te dé los tuits recientes, y que no te devuelva tuits repetidos que ya hayas descargado en una búsqueda anterior. Esto se logra pasando el parámetro **since_id** conteniendo el ids número de algún tuit en particular (digamos 10000000, id número diez millones). Entonces el API de tuiters te devolverá tuits que tengan id mayores a diez millones y tu tendrás confianza que son tuits más recientes al tuit indicado.

Ya que no se puede descargar del API de tuiters muchos miles de tuits de un solo cochacho, puedes ir descargándolos en grupos, y usar el parámetro **since_id** con límite inferior del siguiente grupo de tuits a descargar.