

# Le projet Twitter

Romain Tavenard

## 1 Dates importantes

- Rendu intermédiaire : au plus tard le 30/11/18, 23h59
- Rendu final : au plus tard le 21/12/18, 23h59

## 2 Énoncé

Dans ce projet, vous allez permettre à un utilisateur Twitter d'afficher tous les Tweets de son fil d'actualités contenant un certain mot-clé. Pour ce faire, l'utilisateur devra entrer son `user_id` ainsi que le mot-clé à rechercher et le programme affichera, dans l'ordre chronologique inverse (du plus récent au plus ancien), la liste des Tweets émis ou re-tweetés par un utilisateur qu'il/elle suit. En pratique, on n'affichera pas tous les Tweets du fil d'actualité contenant le mot-clé, mais seulement ceux datant, au plus, d'un certain nombre de jours `max_jours` lui aussi spécifié par l'utilisateur.

## 3 Données à votre disposition

### 3.1 Données à votre disposition (étape 1)

Dans un premier temps (et jusqu'au rendu intermédiaire), votre programme prendra en entrée 3 fichiers :

- a. Un fichier `followers.txt` de la forme :

```
jojo : michel, jean_paul
michel : micheline
[...]
```

Dans l'exemple ci-dessus, l'utilisateur `jojo` est suivi par deux utilisateurs : `michel` et `jean_paul`.

- b. Un fichier `tweets.json` de la forme :

```
{
  "tweet_id1": {
    "date": "10/10/2015, 18:56",
    "user_id": "jojo",
    "text": "blah blah"
  },
  "tweet_id2": {
    "date": "18/02/2017, 19:32",
    "user_id": "micheline",
    "text": "blah blah (bis)"
  },
  [...]
}
```

dans lequel `tweet_id1` et `tweet_id2` sont des identifiants de tweets (**et non pas d'utilisateurs**) pour lesquels trois informations sont accessibles : la date à laquelle ils ont été émis, l'utilisateur qui les a émis et le texte du tweet.

c. Un fichier `retweets.json` de la forme :

```
{
  "tweet_id1": [
    {
      "date": "11/10/2015, 12:34",
      "user_id": "michel"
    },
    {
      "date": "10/10/2015, 19:24",
      "user_id": "micheline"
    }
  ],
  "tweet_id2": [
    {
      "date": "19/02/2017, 07:12",
      "user_id": "jean_paul"
    }
  ],
  [...]
}
```

associant à chaque identifiant de tweet la liste de ses re-tweets (définis par un nom d'utilisateur et une date de re-tweet).

Pour ce premier rendu, **vous devrez vous-mêmes créer des fichiers de test** qui soient pertinents pour le problème étudié et évaluer votre programme sur ces fichiers.

### 3.2 Données à votre disposition (étape 2)

Dans un deuxième temps (rendu numéro 2), ces informations ne vous seront plus fournies sous la forme de fichiers mais seront à récupérer via le module **Tweepy** directement sur Twitter.

Pour cela, le code que vous aurez produit à l'étape 1 sera conservé et seules les instructions de récupération des données devront être modifiées.

Les trois comptes Twitter à considérer sont les suivants :

— @suspect\_robert

- @suspect\_jeanmi
- @HenriSuspect

Et pour tester votre code, vous rechercherez les tweets :

- datant de moins de 50 jours
- visibles par @suspect\_robert
- contenant le mot-clé “coupable”

Comme indiqué en CM, vous ne travaillerez pas directement sur la timeline d’un utilisateur mais sur la liste de ses tweets : vous devrez donc créer une fonction qui récupère, pour un utilisateur donné, ses tweets (et non pas tous les tweets de sa timeline).