**Data Collection Documentation**

**Motivation :**

This dataset (consisting of the hashtags GBV, rape, sexism) collected for analysis and action could mitigate the risk and impact of GBV in Africa. Indeed, all of these tweets are likely to be based on the experiences of those who have experienced violence or those around them, and therefore encourage people not only to be much more cautious but also to take drastic safety measures. This will help to reduce the risk.

**Setps of collection :**

In this competition we scrapped our dataset from Twitter :

The steps required to scrape data from Twitter are :

1.We had usedthe **tweepy** Python library

2.We had applied for a developer account to access the API

3.We had a set of keys to connect to the Twitter API. Specifically, you’ll have an:

* API key
* API secret key
* Access token
* Access token secret

4. These keys are inserted directly into our code to connect and to the Twitter API.

5.We add hachtags related to gender-based violence ('#GBV','#sexism','#rape' )

5.We fetched all tweets that have specific hashtag

6. We converted our tweets into DataFrame

7. We saved the dataframe into cs