

# Homework Assignment 1

Due November 17<sup>th</sup>, 1pm

# Homework: 'UNGA\_SPEECHES.csv'

Using what we have learned during the class (NLP, text analysis, wordcloud):

1. Find top 10 most common nouns of all speeches within the file.  
(top 10 most common nouns of the **entire csv file**, not individual countries)
2. Make **word cloud** for **four country of your choice** based on word frequency.
  - If you want to specifically add countries, you can do that.
  - Just add them to your csv file before importing.
3. Measure **polarity & subjectivity** scores for each country using TextBlob.
4. Hand in two files by the deadline (November 17<sup>th</sup>, 1pm.)
  - a. the ipynb file
  - b. a word file with your four word clouds

## Tips:

- Try to refine them as much as possible (e.g: lemmatization, include stopwords, take out stopwords if you regard it unnecessary etc)
- You can decorate your word cloud (giving shapes and colors)

# Your final product to be submitted

Your name

Word cloud of country A

polarity score: # , subjectivity score: #

Word cloud of country B

polarity score: # , subjectivity score: #

Word cloud of country C

polarity score: # , subjectivity score: #

Word cloud of country D

polarity score: # , subjectivity score: #