**Angelo Basile**

s3275655
a.basile@student.rug.nl

# Personality Prediction

**30th April 2017**

## OVERVIEW

The work of [Liu et al., (2016)] shows that neural models can successfully model author personality traits from text; tweets in that case. The authors suggest at the end of the paper that "the lack of feature engineering should support language independence". With this project I will make use of the TwiSty [Verhoeven et al., 2016] corpus to test that hypothesis.

## GOALS

1. Predicting author personality traits using a single model for different language
2. Comparing a neural model against a BOW/SVM model baseline.

## SPECIFICATIONS

The TwiSty corpus is a multi-lingual (Dutch, German, French, Italian, Portuguese and Spanish). tweet corpus annotated with MBTI type indicators. There are 16 labels. The modelling will replicate the one described in [1]. # ToDo: add details on method later as course goes on.

## PLAN

1. Describe distribution of labels in the corpus
2. Run baseline model
3. Explore neural model

## BIBLIOGRAPHY

[1] Liu et al., (2016) http://aclweb.org/anthology/W/W16/W16-4303.pdf

[2] Verhoeven et. al, (2016), TwiSty: a multilingual twitter stylometry corpus for gender and personality profiling. In Proc. LREC, pages 1632–1637, Portoroˇz, Slovenia.