**Documentation:**

Configuration files and methods for pre- and post-processing datasets

# 15-12-2021

****

# Configuration json file

The configuration json file contains all the rules for fixing errors or applying specific definitions that a medical terminology may require. Configuration rules may be applied before the training process (pre-processing the training dataset) or after inference (post-processing the output translation).

# Configuration script file

The configuration scripts files contains all the methods required for the pipeline to apply the configuration rules, as stated the json file.

um8\_unit\_replace(string):

"""Rule um8

The letter μ meaning micro in a unit is replaced by the letter u.

"""

check\_for\_mg(string):

check for measurement errors