LINKS:

Dataset: <https://github.com/roneysco/Fake.br-Corpus>

Pesquisa USP do Dataset: <https://sites.icmc.usp.br/taspardo/PROPOR2018-MonteiroEtAl.pdf>

Video legal webscraping: <https://www.youtube.com/watch?v=Vxl5jUltHBo>

Pra rodar selenium tem que baixar isso <https://github.com/mozilla/geckodriver/releases>



Conteúdos legais a serem extraídos do texto origem para serem pesquisado nos sites de confiança:  
 - conteúdo mais relevante (word frequence)  
 - local  
 - data

- personalidades famosas

Identificar a sequência de palavras no primeiro parágrafo do texto das notícias dos grandes portais.

Sites confiáveis:

G1