

## Rezumat / Abstract

Lucrarea intitulată „CoRoSeOf – An Annotated Corpus of Romanian Sexist and Offensive Tweets” prezintă **CoRoSeOf**, un corpus mare de social media în limba română, adnotat manual pentru limbaj sexist și ofensiv. Descriem procesul de adnotare a corpusului, oferim analize inițiale și rezultatele de clasificare de bază pentru detectarea sexismului pe acest set de date. Corpusul rezultat conține 39 245 de tweet-uri, adnotate de mai mulți adnotatori (cu o rată de acord de  $\kappa$  de Fleiss = 0,45), urmând setul de etichete sexiste dintr-un studiu recent. Detectarea automată a sexismului produce scoruri similare cu unele dintre studiile anterioare (scorul F1 macro mediu de 83,07 % în sarcina de clasificare binară). Eliberăm corpus-ul cu o licență permisivă.

The paper titled „CoRoSeOf – An Annotated Corpus of Romanian Sexist and Offensive Tweets” introduces **CoRoSeOf**, a large corpus of Romanian social media manually annotated for sexist and offensive language. We describe the annotation process of the corpus, provide initial analyses, and baseline classification results for sexism detection on this data set. The resulting corpus contains 39 245 tweets, annotated by multiple annotators (with an agreement rate of Fleiss'  $\kappa$  = 0.45), following the sexist label set of a recent study. The automatic sexism detection yields scores similar to some of the earlier studies (macro averaged F1 score of 83.07 % on binary classification task). We release the corpus with a permissive license.