



1



2



3

## EXPLORATORY ARABIC OFFENSIVE LANGUAGE DATASET ANALYSIS

Fatemah Husain

[f.husain@ku.edu.kw](mailto:f.husain@ku.edu.kw)

Ozlem Uzuner

[ouzuner@gmu.edu](mailto:ouzuner@gmu.edu)

### Abstract

This paper adding more insights towards resources and datasets used in Arabic offensive language research. The main goal of this paper is to guide researchers in Arabic offensive language in selecting appropriate datasets based on their content, and in creating new Arabic offensive language resources to support and complement the available ones.

### Introduction

Annotated offensive language datasets are used to categorize texts according to their offensive content automatically. As it is mentioned previously, some examples of offensive content are hate speech, obscene language, or vulgar language. The automated categorization