

Abstract

Radu Meza

**A Digital Social Science Approach
to Studying Hate Speech on
Facebook in Romania and Hungary**

Online hate speech is the subject of heated policy debate in Europe and globally. Although hate speech acts have been regulated by European laws, online communication through platforms owned by businesses located outside the users' country and subject to different legislation raises new issues. The 2015 UNESCO study on "Countering Online Hate Speech" identifies definition, jurisdiction, comprehension and intervention as key aspects. When analyzing hate speech as an act of communication, researchers have focused on: content (what is being said), emitters (who is communicating), targets (who is the message about), and context (including when the act takes place). From a methodological standpoint, detecting violent, obscene or hate speech is a problem for both media researchers and content managers or digital platform owners. Natural language processing is a complex task and there is a scarcity of tools available for most languages. Recent efforts in computational linguistics use machine learning techniques similar to sentiment analysis in correlation with techniques for detecting terms used online to reference racial, ethnic or religious groups. The proposed paper aims to use digital social research methodology to explore the issue of hate speech on public Facebook contexts through a comparative analysis of a purposeful sample of several hundred thousand comments posted to twenty Romanian and Hungarian

Facebook Fan pages (news outlets, satirical pages, political parties/leaders) in 2016 and 2017 to identify frequent targets and attempt a classification of content.