

Towards an Early Warning System for Online and Offline Violence

Diana Maynard

University of Sheffield

Gender-based online violence against women journalists is one of the biggest contemporary threats to press freedom globally. This talk describes a dashboard we are developing for monitoring and exploring relevant social media data, as well as some findings in the form of recently published big data case studies investigating online violence targeted at a number of emblematic women journalists from around the world. In order to conduct this large scale analysis of online abuse, we have developed NLP tools to identify and characterise online abuse from Twitter targeted at specific individuals, with the ultimate aim of developing an early warning system to help predict the escalation of online abuse into offline harm and violence, based on indicators from the analysis. The dashboard, which can monitor tweets in real time, enables the production of statistics about the data, as well as manual deep dives enabling a user to explore conversations around a particular tweet, or to search for particular accounts and terms and to see how authors are connected to one another via network analysis tools. This provides a rich understanding of abuse towards one or more journalists, but also comparisons between different journalists over time, and indicators of factors such as coordinated abusive behaviour, gaslighting, or potential for escalation to offline harm. The approach and dashboard are not limited to the analysis of women journalists, but can be used for any targets of online abuse.