Abstract
This paper describes an approach taken to analyze and categorize a sizable dataset of politically oriented posts that were submitted to a popular idea bank, Egypt 2.0, created following the Egyptian revolution. The aim of the analysis was to organize and present the data in a simple way that allows the voice of the people to be heard by decision makers and activists in a critical 6 week period in February and March 2011. The constraints faced when developing the approach included the absence of a classification scheme, the unavailability of training data, the need to assign more than one category, or label, to individual posts and the need to complete the task in a short period of time. The goal of this paper is twofold. Firstly, to present and evaluate the rapid development framework and algorithms used to organize the data. Secondly, to document the challenges encountered when both developing the system itself and analyzing the data, and to present our experience to the research community with the aim of identifying potentially new interesting research topics.

References

What the Masses Want: A Case Study in Knowledge Discovery from Politically Oriented Data


Index Terms

Computer Science

Information Sciences

Keywords

Text Mining; Text Analysis; Topic Categorization; Multi-Labeling