Quantifying Politics Using Online Data: Introduction

Yelena Mejova[1] and Ingmar Weber[2]

Large web-based data sets make possible political studies at a scale inconceivable just a decade ago. Everything from personal opinions to popular political movements leaves a footprint online and provides a firsthand account of both everyday and historic events. These data are, however, unstructured and full of noise requiring new approaches—quantitative methods developed in the realms of political and social science, but also in data analysis and mining. Applied to online data, these make possible language modeling, topic tracking, novelty detection, social network mining, and many more types of analyses, all providing new insights into social and political realities.

The works compiled below are a sample of approaches to quantify political phenomena using online data. All articles underwent a peer-review process by experts from both computer science and/or political/social science. The topics covered include general-purpose social media like Twitter (Crooks et al., Dyagilev et al.) and blog posts (Sagi & Dehghani), news outlets (Yom-Tov et al., Bright & Nicholls, Sagi & Dehghani), search query logs (Yom-Tov et al.), governmental websites (regulations.gov in Levy & Franklin), and political citizen-driven sites (debatepolitics.com in Liang). The findings in these works illustrate the richness of the data available online, and the tools one may employ to quantitatively assess political opinions expressed there, the use and readership of online content, and the various social networks people form, as witnessed through their actions online.

We believe that these scientific articles present timely research in an area that will continue to grow and draw more attention. Hopefully, this special issue can contribute to make this research area more interdisciplinary and initiate new discussions and collaborations.

Author Biographies

Yelena Mejova (ymejova@yahoo-inc.com) is a postdoctoral researcher at Yahoo! Research in Barcelona, Spain. Specializing in text retrieval and mining, she created and analyzed multiple web-based data sets, including webpages, blogs, reviews, and Twitter. This analysis included sentiment detection, political opinion extraction, and topic tracking, and in particular the political support classification and evaluation.

Ingmar Weber (ingmarweber@acm.org) is a senior scientist at Qatar Computing Research Institute. His research covers a wide subject area from classical information retrieval, to sponsored search, with recent work focusing on computational political science and interdisciplinary studies in web science. He has studied the polarization in U.S. politics in web search and on Twitter, and is currently investigating Arab politics in social media.

[1] Yahoo! Labs, Barcelona, Spain
[2] Qatar Computing Research Institute, Doha, Qatar

Corresponding Author:
Yelena Mejova, Yahoo! Labs, Avinguda Diagonal 177, Barcelona 08018, Spain.
Email: ymejova@yahoo-inc.com