Svetlana S. Bodrunova

St. Petersburg State University

[s.bodrunova@spbu.ru](mailto:asdfg@ghj.ru)

**What is opinion online? Automated opinion mining and its current limitations**

Today, the growing field of opinion mining needs systematization that should start from (re)conceptualizing of what an opinion is. Today, we see a major cleavage between how opinions are conceptualized in ‘traditional’ theory-driven science and data-driven science based mostly on automated methods of big textual data analysis. We review the relevant theory-driven approaches that add to automated opinion mining and show that data-driven science is largely limited by binary ‘yes/no’ views upon user opinions, excessive orientation to social network analysis, disregard of the differences between the macro- and micro-levels of opinion formation, and, most important, lack of reflection on the gap between AI-based and human understanding of what opinion is.

Keywords: public opinion, opinion mining, big data analysis, cumulative opinion formation, opinion patterns

Recently, public opinion research has been expanding; it has already included opinions expressed via new forms of mediatized (especially socially-networked) communication. Computational studies have provided for new approaches in opinion mining and detection different from traditional polling, and several new approaches to what ‘opinions’ are have emerged.

Even before the advent of online social networking, there was a lack of all-encompassing and multi-faceted theory of opinion. In practical sociology, polling methods and technologies were highly developed but could not grasp the day-by-day opinion dynamics and saw opinions as attitudes, often of close-to-the-‘yes/no’ nature. In linguistics, opinions derive from the logic of argumentation; in journalism, opinions possess the forms of whole genres. One of the fruitful attempts to suggest an operational definition of opinion based on social exclusion belongs to Elizabeth Noelle-Neumann [2]; it has shown its applicability to automated and AI-based textual analysis. However, there is no agreement between various branches of social science and humanities on what opinions actually are.

This diversity of approaches to opinion, including those in automated and AI-based opinion detection, has already turned from a spurring advantage to an obstacle for public opinion studies [1]. It also produces smaller-scale conceptual issues that we try to address.

These issues, i.a., are the following. First, the widening differences between the pre-computational and computational opinion studies demand their closer exchange and co-operation. Second, the common conceptual basis of data-driven research often reduces ‘opinion’ to the lexical-semantic levels of speech, largely ignoring social-group, political, cultural, temporal, and narrative aspects of opinion formation. Third, it also reduces opinion detection goals to ‘yes/no’ detection, simplifies user sentiment, or mixes topicality and opinion. Fourth, automated and computational opinion mining is much oriented to the principles of networked communication and large-scale datasets, deprivileging systemic approaches to opinion dynamics and micro-patterns of opinion formation. Fifth, conceptual discussion on whether the gap between human vs. AI-created human-like representation of opinions can ever be overcome is very limited. If we cannot overcome it, we need to pose the fundamental question on how the machine sees opinions, why they cannot reach the logic of human judgment on opinionated content, and, if so, what could be the foundations of our trust to AI-detected opinion.

We suggest to bridge the divisions in conceptualization of opinions created by the (very diverse) ‘traditional’, computational, and human-like computational views on opinions in today’s online discussions. Our main focus is on how the pre-computational and computational methods of opinion detection could enrich each other and whether full similarity between opinion formulation and/or representation by artificial intelligence and human beings could ever be reached. In the conclusion, we hint on the future directions of conceptual research on the nature and essence of human and human-like opinion.

References

1. Bodrunova, S. S. Opinion types on social media: A review of approaches to what opinions are in social vs. computational science. In: Social Computing and Social Media. Design, Ethics, User Behavior, and Social Network Analysis. HCII 2024. Lecture Notes in Computer Science series (in print). Cham: Springer.

2. Noelle‐Neumann, E. (1974). The spiral of silence a theory of public opinion. Journal of Communication, 24(2), 43–51.