The epistemological impacts of big data on public opinion studies

Pedro Caldas, Vinicius Romanini

School of Communication and Arts, University of São Paulo, Brazil.

Abstract

In this work, we seek to highlight and describe the main differences between traditional public opinion polls (made by using methods and techniques traditionally undertaken in the social sciences), and those accomplished through methodological processes made possible by the adoption of big data. We ensure a special focus on the consequences brought about by the use of nonparametric analysis over parametric analysis to show how big data is impacting not only the methodological aspects but the epistemological basis of public opinion studies in general. Researchers see an epistemological struggle between methodology and theory in public opinion studies. This struggle is composed of two approaches: a quantitative one and a qualitative one. On the one hand, we have quantitative polls methods which lead to an excessively contextual representation of public opinion. On the other hand, we have general theories that do grasp public opinion in most of its complexity but fall short in providing sophisticated empirical tools for contextual analysis of public opinion specific issues. The methods undertook by pollsters, as many others used in social sciences rely upon classical scientific structures, where researchers conduct their studies through hierarchical theories and survey techniques to access and understand their subject. In these cases, the researchers must pose the research problem a prioristically, to parametrize and create the questionnaires before the collecting of the data to be analyzed after. By using big data models, the need for posing a research problem and parametrize the proceedings of the study a prioristically no longer exists, thus contributing to a characterization of public opinion that is qualitative and way more complex, rather than the traditional one. Although not yet strictly statistically representative, public opinion studies made by using datasets collected from social media provide us with a view of public opinion that shows, among other things, the main actors (persons, groups, and organizations), their powers of influence over the others and their interests in public opinion formation movement.

Keywords: Public opinion; big data; polls; mining; epistemological impacts.