

Data and Discourse around Europe

Wolf J. Schünemann

Datafication has reached political science, at least through the interdisciplinary backdoor called Computational Social Science (CSS). More and more, political scientists across the sub-disciplines feel encouraged to overcome the micro-orientation and methodological individualism of mainstream political studies by turning towards digital trace data and computational linguistics – which is basically good news. However, they are somehow challenged by computer scientists and computer linguists entering the field with a combination of higher technical proficiency and lesser theoretical reflexivity. At the same time, there is a tendency within the discipline to perceive CSS approaches and methods as innovative developments in quantitative research methods only, which leads to new alliances based on shared data-driven ideals of political research.

The basic argument of the talk is that this latter tendencies are problematic as they are based on fundamental misperceptions. First, meaningful studies still seem not to work in a purely data-driven manner but need to be theory-driven. Moreover, assessments in CSS studies are mostly not made merely based on quantitative measurements but they do include qualitative work that has to be made transparent. This qualitative work requires subject-related knowledge and theoretical premises.

The more general questions about data science are illustrated by revisiting a comparative discourse study on EU referendum campaigns in the years 2000. The referendums were part of the so-called constitutional process, in which the electorates of France and the Netherlands voted down the Constitutional Treaty in 2005. Three years later, the less symbolic Lisbon Treaty failed in a first Irish referendum and was only adopted in a second referendum a year later. The experiment presented in the talk is based on the qualitative analysis of the comprehensive corpora. Different tools from computer linguistics and CSS, in particular topic modeling, are tested and evaluated depending on how they help to answer core questions of EU referendum research. In the end, Structural Topic Modeling (STM) is presented and discussed as a potential means for qualitative discourse analysis in EU studies.