Cross-lingual Classification and Scaling of Political Texts

Political text scaling aims to linearly order parties and politicians across political dimensions (e.g., left-to-right ideology) based on textual content (e.g., politician speeches or party manifestos). Existing models, such as Wordscores and Wordfish, scale texts based on relative word usage; by doing so, they do not take into consideration topical information and cannot be used for cross-lingual analyses. In my talk, I'll present our efforts towards developing a topic-based and cross-lingual political text scaling approach. First we'll introduce our initial work, TopFish, a multi-level computational method that integrates topic detection and political scaling and shows its applicability for temporal aspect analyses of political campaigns (pre-primary elections, primary elections, and general elections). Next, we'll present the work on cross-lingual topic-classification. Finally, we'll show a new text scaling approach that leverages semantic representations of text and is applicable to cross-lingual political text scaling.