Extracting Socioeconomic Patterns from the News: Modelling Text and Outlet Importance Jointly

**DATA**
- News summaries written in English
- Extracted from the Open Europe Think Tank
- Daily aggregation of news about the EU
- Focused on current affairs, politics, economy
- Each summary is a few paragraphs long
- Each summary is attached to ≥1 news outlets
- 1913 days, 94 months
- Avg. 14 news/summary
- Features: 8413 unigrams + 19045 bigrams
- 435 news outlets

*ESI* = Economic sentiment indicator
- Composite indicator often seen as an early predictor for future economic developments
- Consists of five confidence indicators with different weights: industrial (40%), services (30%), consumer (20%), construction (5%) and retail trade (5%)
- Any socioeconomic indicator can be used in this framework