How to identify 'umbrella' concepts not spoken out? Exploring German and Finnish plenary debates on 'Democracy' (1990-2020) with a TNA-based method

Dr Kimmo Elo
University of Turku, Finland, Centre for Parliamentary Studies
E-mail: kimmo.elo@utu.fi

Proposal for oral presentation on the '3rd Conference on Digital Data and Human Sciences (DRDHum 2024)', Dec 2024, University of Eastern Finland

A very common way to analyse issue-specific political debates is to use keyword-based queries to identify relevant documents. Although proven to provide a rather solid material base for analysis, when measured against known contextual attributes like temporal distribution, the representativeness of the sample is far more difficult to be evaluated. With very concrete and unambiguous concepts (e.g. climate change, debt) a standard word-based query most probably returns reliable results.

But how about 'umbrella' concepts, whose definition often integrates a bunch of lower-level, yet powerful and 'loaded' concepts and which are used as discursive frames even if not verbally mentioned? For example, someone can speak about costs and revenues, products and competition so that it becomes clear that she also talks about 'market economy', even if she never say 'market economy'.

My paper will introduce a powerful, yet easy-to-use, word list (a.k.a. dictionary) based method for the identification of discursive frames from a document corpus. The method itself integrates Text Network Analysis (TNA)

with both *tf-idf* and collocational analysis in a snowballing method to explore contextual frames. The tool allows a researcher to explore and browse her document corpus as a clustered concept collocation network, in which a cluster represents a discursive frame based on concept collocations. As the concept network can be enriched e.g. with time-related information, it allows a researcher to explore how discursive frames surface, evolve and vanish in a certain period of time. Being wholly aware of quantum leaps made in both semantic and AI based search engines, my approach defends is usefulness and relevance by integrating the power of computational methods with a corpus- and context-aware approach to textual data.

My paper will exemplify and evidence the power of the underlying method through a comparative analysis of Finnish and German plenary debates on 'democracy' from 1990 to 2020.

Keywords: Plenary debates, Exploratory Data Analysis, Data mining, Conceptual history, Text Network Analysis, Democracy, Finland, Germany