

WP 3.5 TEXT NETWORK ANALYSIS OF POLITICAL TEXTS

Introduction

The WP 3.5 will establish a research infrastructure to structure, store, and analyse large textual document corpora. The first data to be integrated in this infrastructure consists of main political documents of the Finnish politics, including all parliamentary minutes from 1907 onwards, government programmes, as well as party programmes.

The method toolbox provided by the WP 3.5 will consist of a variety of tools for Text Network Analysis (TNA). The tools will be available online, as well as dedicated R packages for offline use.

This WP is a collaboration project between three research units at the University of Turku: Centre for Parliamentary Studies, TurkuNLP Group, and Turku Data Science Group.

FinParl Corpus

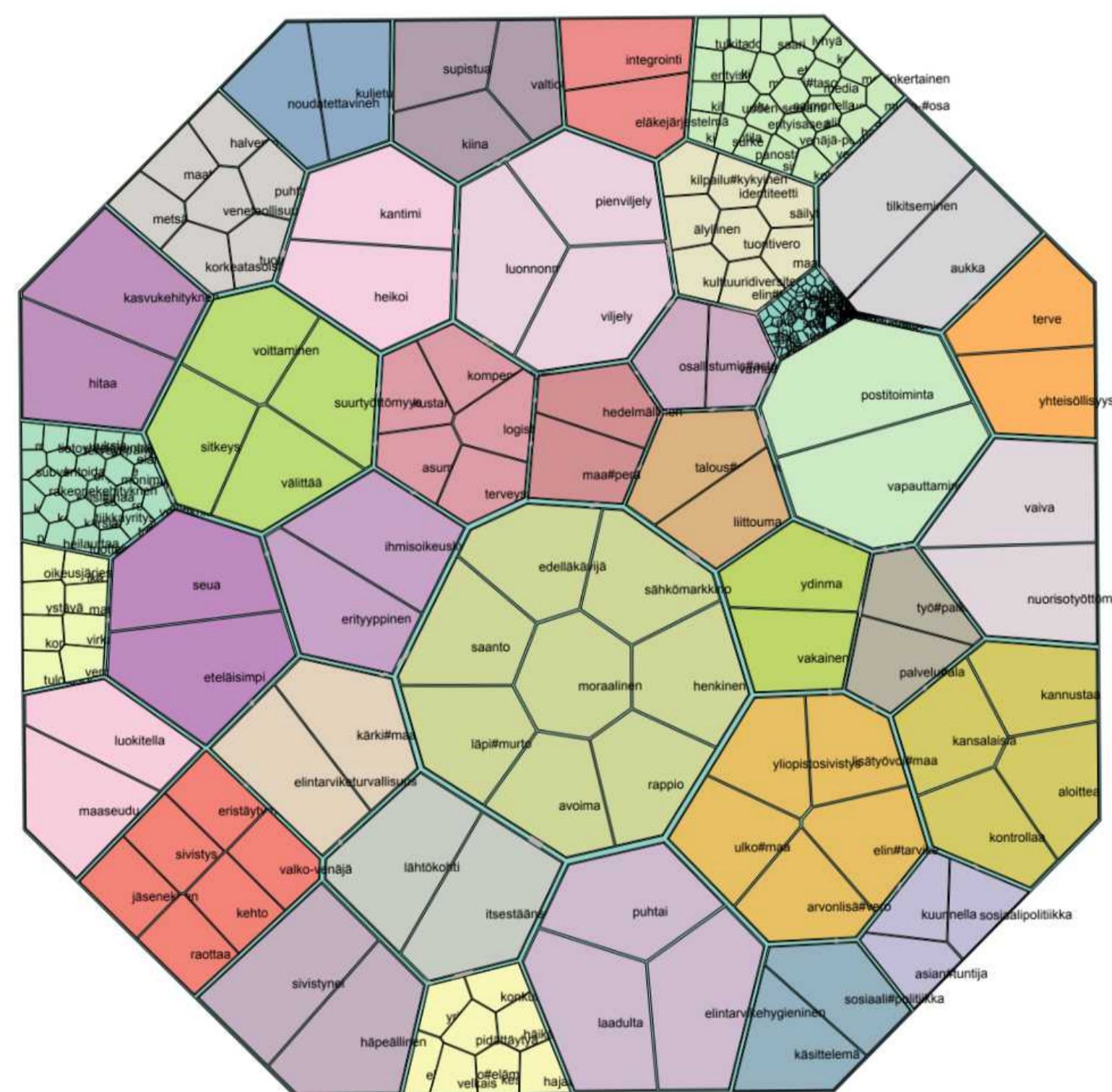
The early-stage development of the WP 3.5 is build on the **FinParl** dataset consisting of plenary minutes of the Finnish Parliament (*Eduskunta*) from 1907 to 2021. The corpus has been created as part of the “Semantic Parliaments” research consortium project financed by the Academy of Finland (2020-2022). This consortium project is lead by the *Semantic Computing Research Group* of the Aalto University, other participants are the *HELDIG Centre* of the University of Helsinki, and the *Centre for Parliamentary Studies* at the University of Turku.

Plenary minutes were digitised and curated into **FinParl** database, a machine-readable XML database. The XML dataset follows the Parla-CLARIN standard, as this is mostly used in similar international projects. The database has a rich set of metadata including e.g. the speaker’s name, possible roles (speaker/deputy speaker, minister, president, etc.), the parliamentary group of the speaker, as well as information about the session, in which the speeches was given[1].

The corpus will be made available as open data in January 2023. The corpus will also be subject to continuous updating after its release.

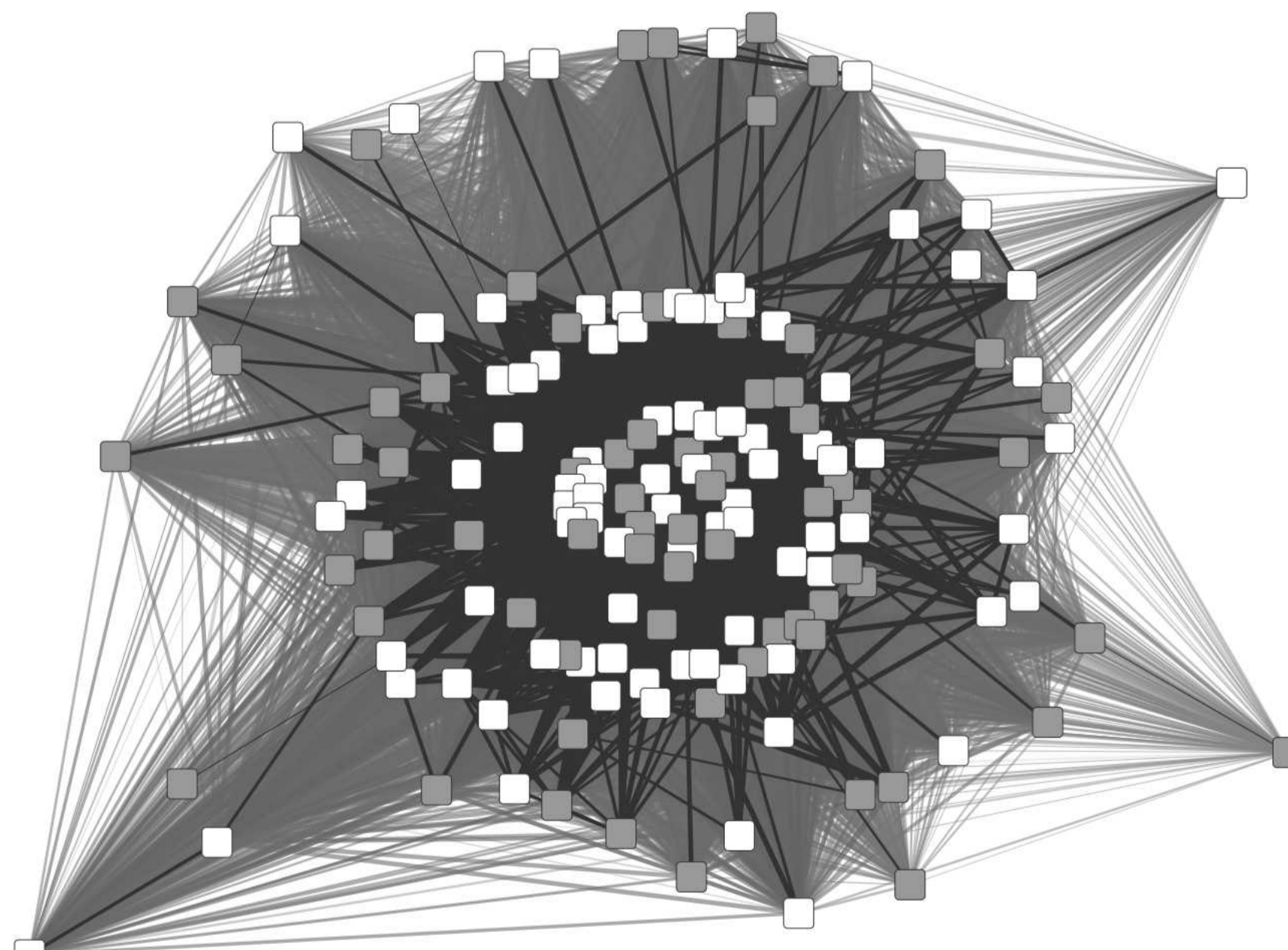
Text Network Analysis (TNA)

TNA as a method is a spin-off of Social Network Analysis (SNA), defining a network as a set of dots (nodes, vertices) and connecting lines (edges) between nodes. TNA can be used e.g. to explore repetitive patterns hidden in the text’s structure, to identify salient concepts for the study of meanings and frames, or to analyse similarities and differences between different speakers and over time.



Thematic clusters in a text network of plenary debates on European integration in the Finnish Eduskunta 1990-2020 (layout: Voronoi Treemap)[2].

The true analytical power of TNA lies in the possibility to apply robust methods and techniques of network analysis to the text network. These tools include a multiple of visualisation techniques, tools for the analysis of network structure, as well as methods for assessing a word’s or concept’s status and role within a network. Further, TNA can rather easily be used with different theories in order to explain the findings based on TNA.



Cosine distance network of Finnish MPs during the plenary debate on Finland’s application to NATO on May 16-17, 2022 (layout: Backbone).[3]

TNA for web data?

Web data offers fruitful possibilities for analysing the evolution and circulation of ideas, discussion topics and opinions. TNA can provide an efficient method for exploring this.

Contact

Dr. Kimmo Elo

Centre for Parliamentary Studies

University of Turku

Email: kimmo.elo@utu.fi

Phone: +358 29 450 4683

WWW:

<https://www.utu.fi/en/university/faculty-of-social-sciences/centre-for-parliamentary-studies>

Prof. Veronika Laippala

School of Languages and Translation Studies

University of Turku

Email: veronika.laippala@utu.fi

WWW: <https://turkunlp.org/>

- [1] L. Sinikallio, S. Drobac, Minna Tamper, Rafael Leal, Mikko Koho, Jouni Tuominen, Matti La Mela and Eero Hyvönen (2021): Plenary Debates of the Parliament of Finland as Linked Open Data and in Parla-CLARIN Markup. *Open Access Series in Informatics (OASICs)*, 93(8), pp. 1–17. <https://drops.dagstuhl.de/opus/volltexte/2021/14544/pdf/OASICs-LDK-2021-8.pdf>
- [2] Elo, Kimmo (2022): Debates on European Integration in the Finnish Parliament (*Eduskunta*) 1990-2020. In: Matti La Mela, Fredrik Norén, Eero Hyvönen (eds.): *Proceedings of the Digital Parliamentary Data in Action (DiPaDA 2022) Workshop*, CEUR vol. 3133, pp. 129–145, <http://ceur-ws.org/Vol-3133/paper09.pdf>
- [3] Elo, Kimmo (2022): *Data-analyttisiä koeporauksia Suomen eduskunnan Nato-keskusteluun 16.-17.5.2022*. Research Note, Turku Centre for Parliamentary Studies, <http://dx.doi.org/10.13140/RG.2.2.20871.68004>