Aalto University
School of Science

UNIVERSITY OF HELSINKI

SeCo

HELDIG
Helsinki Centre for Digital Humanities

Analyses of Networks of Politicians Based on Linked Data: Case ParliamentSampo – Parliament of Finland on the Semantic Web

SWODCH 2022
Turin, 2022, 5 September
Henna Poikkimäki, Petri Leskinen, Minna Tamper, Eero Hyvönen

Contents

- *ParliamentSampo* **project**

Contents

- *ParliamentSampo* **project**
- **Named entity recognition and linking**

Contents

- *ParliamentSampo* **project**
- **Named entity recognition and linking**
- **Network analysis**

ParliamentSampo Project

Aalto University
School of Science

UNIVERSITY OF HELSINKI

SeCo

HELDIG
Helsinki Centre for Digital Humanities

Project Consortium

- **University of Helsinki, Helsinki Centre for Digital Humanities (HELDIG)**
  - Consortium coordinator
  - Main focus: Language analysis and technologies
  - Eero Hyvönen (PI)
- **Aalto University, Department of Computer Science, Semantic Computing Research Group (SeCo)**
  - Main focus: Linked Data and Semantic Web technologies, AI
  - Data services and infrastructures
  - Jouni Tuominen (PI)
- **University of Turku, Centre for Parliamentary Studies**
  - Main focus: Political and Media Studies
  - Kimmo Elo (PI)

The project is
funded by **the Academy of Finland**

Project *ParliamentSampo*

- Main contributions
  - Finnish parliamentary data as a national Linked Open Data (LOD) infrastructure and service

Project *ParliamentSampo*

- Main contributions
  - Finnish parliamentary data as a national Linked Open Data (LOD) infrastructure and service
  - Studies of political culture and language

Project *ParliamentSampo*

- Main contributions
  - Finnish parliamentary data as a national Linked Open Data (LOD) infrastructure and service
  - Studies of political culture and language
  - Enrich semantically content in other related Finnish LOD services

Project *ParliamentSampo*

- Main contributions
  - Finnish parliamentary data as a national Linked Open Data (LOD) infrastructure and service
  - Studies of political culture and language
  - Enrich semantically content in other related Finnish LOD services
- Related to similar efforts e.g. in Italy and Latvia and in European Union

Project *ParliamentSampo*

Data publication consists of two interlinked knowledge graphs

1) **Speeches**
   Parliamentary debate speeches
2) **Actors**
   Members of Parliament of Finland, groups and organizations

Named Entity Recognition and Linking

Named Entity Recognition (NER)

1. **Query textual speeches from SPARQL endpoint**

Named Entity Recognition (NER)

1. **Query textual speeches from SPARQL endpoint**
2. **Clean speeches for interruptions and lemmatize**

Named Entity Recognition (NER)

1. **Query textual speeches from SPARQL endpoint**
2. **Clean speeches for interruptions and lemmatize**
3. **Use FinBERT-NER model for extracting place, person, organization, expression of time and legislation mentions**

Named Entity Recognition (NER)

1. **Query textual speeches from SPARQL endpoint**
2. **Clean speeches for interruptions and lemmatize**
3. **Use FinBERT-NER model for extracting place, person, organization, expression of time and legislation mentions**

- **For 100 mentions: precision 97%, recall 77%, F1-score 86%**

Named Entity Linking (NEL)

- **Internal linking to ParliamentSampo actor knowledge graph**
- **External linking for broader data enrichment**

Table 1. Metadata schema for the class for NamedEntity.

Element URI	C	Range	Meaning of the value
:surfaceForm	1	xsd:string	Original surface form in text
:count	1	xsd:integer	Number of entity mentions in a speech
:category	1	xsd:string	Type of the named entity
skos:relatedMatch	0..*	rdfs:Resource	Links to ontologies for named entities
provo:wasAssociatedWith	1..*	:NamedEntityMethod, provo:SoftwareAgent	Provenance information about the method used to extract the named entity

Namespace *provo* refers to PROV-O ontology

Named Entity Linking (NEL)

- **For people, only full name mentions were linked successfully**
  - Family name mentions have since been linked

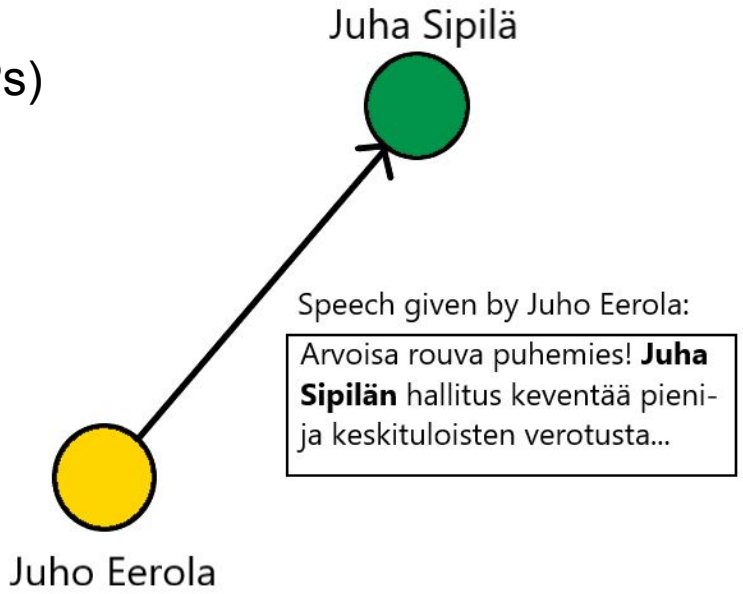
Named Entity Linking (NEL)

- **For people, only full name mentions were linked successfully**
  - Family name mentions have since been linked
- **For 50 speeches with 105 person name mentions: precision 95%, recall 80%, F1-score 87%**

Network Analysis

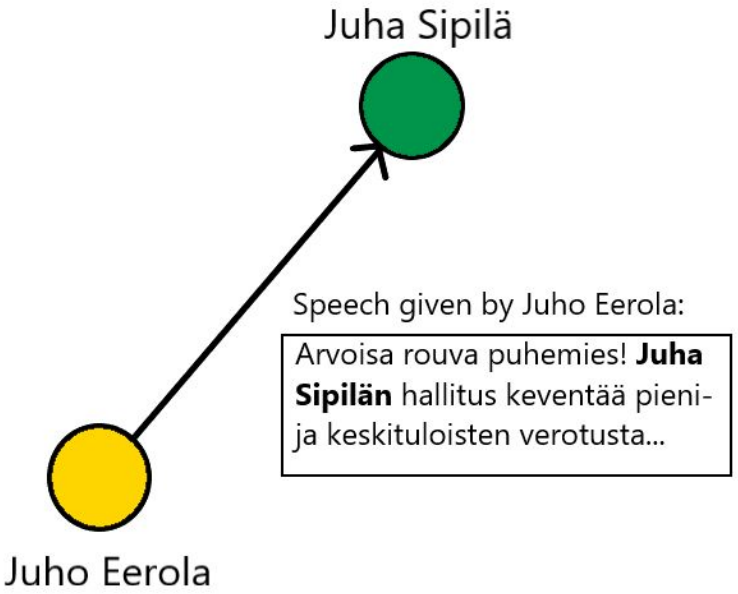
Reference Networks

- Nodes: Members of Parliament (MPs) or parties
- Links: Mentions between MPs or parties



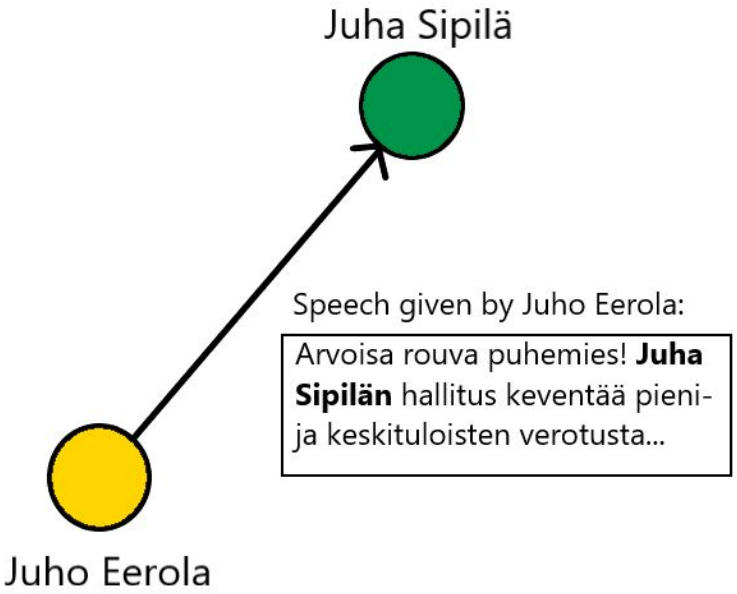
Constructing Reference Networks

1. Query links from chosen subset of speeches

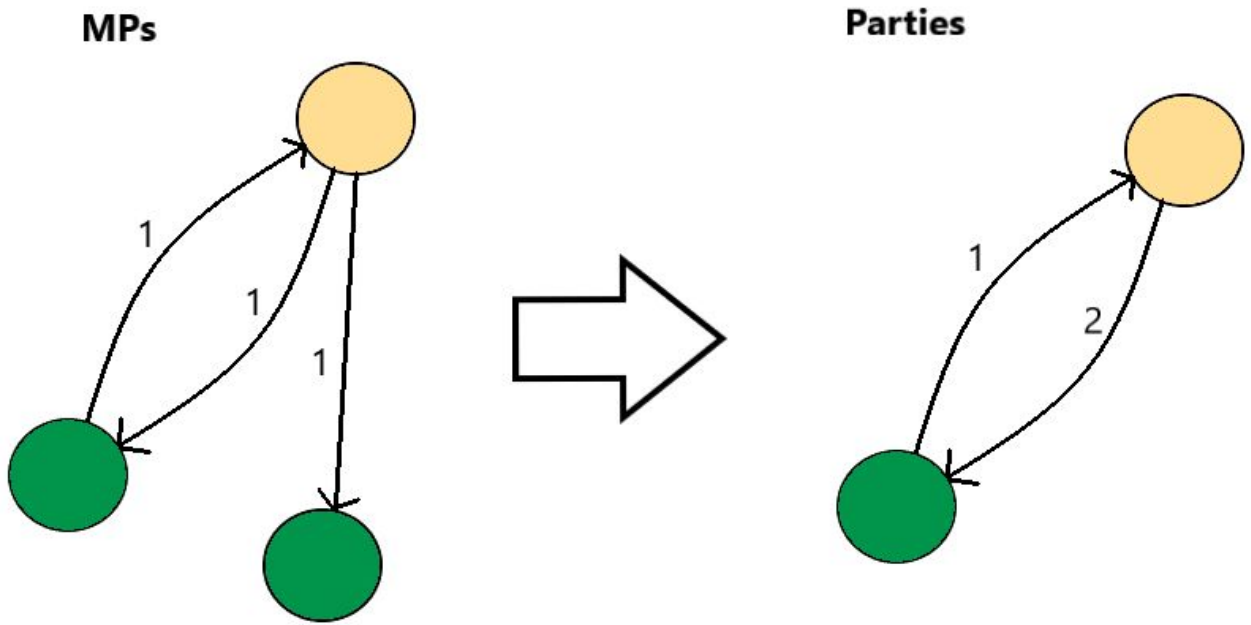


Constructing Reference Networks

1. Query links from chosen subset of speeches
2. Query metadata for nodes

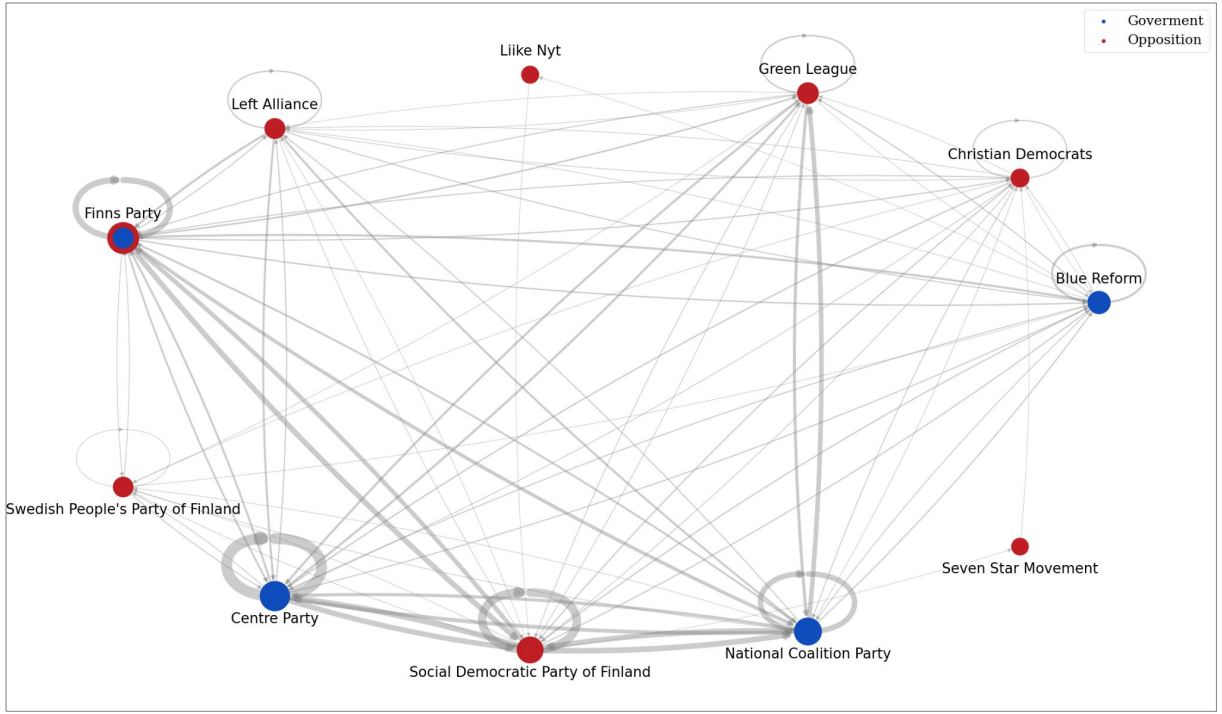


Reference Network for Parties

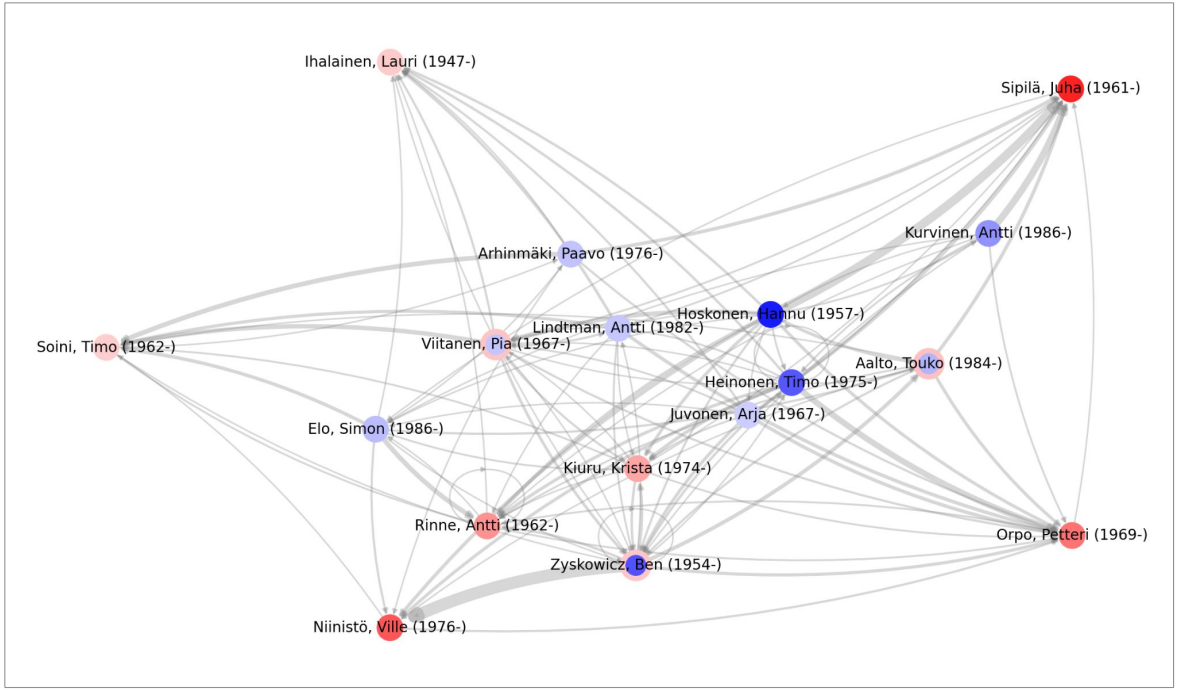


Network Analysis

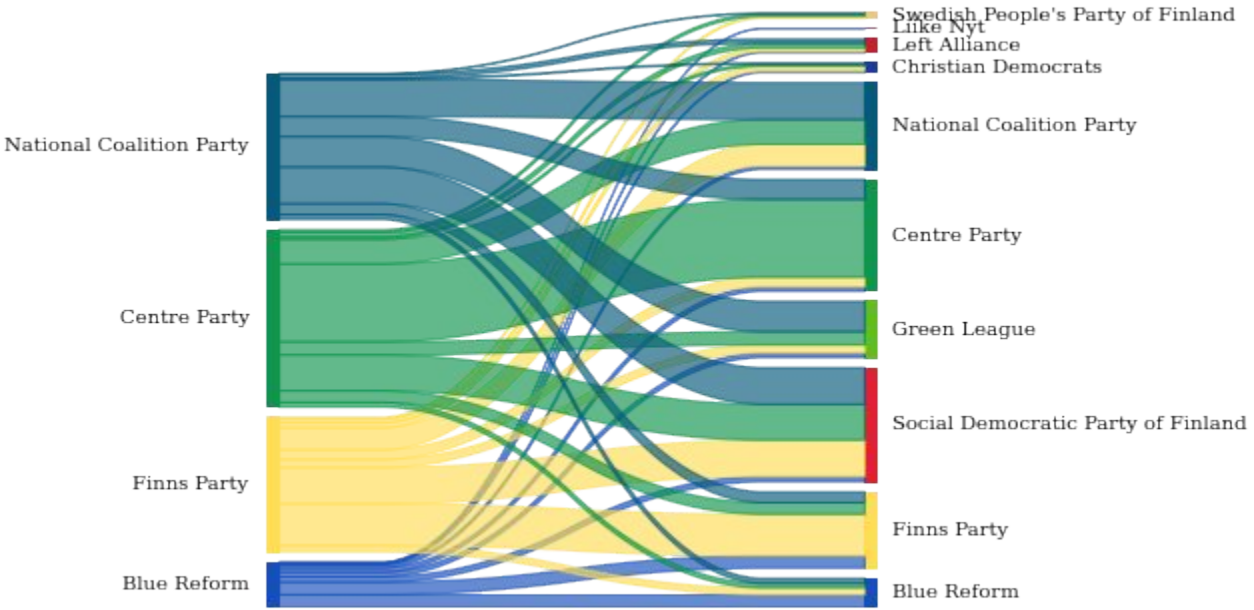
- 209 MPs
- 11 parties
- 2100 mentions



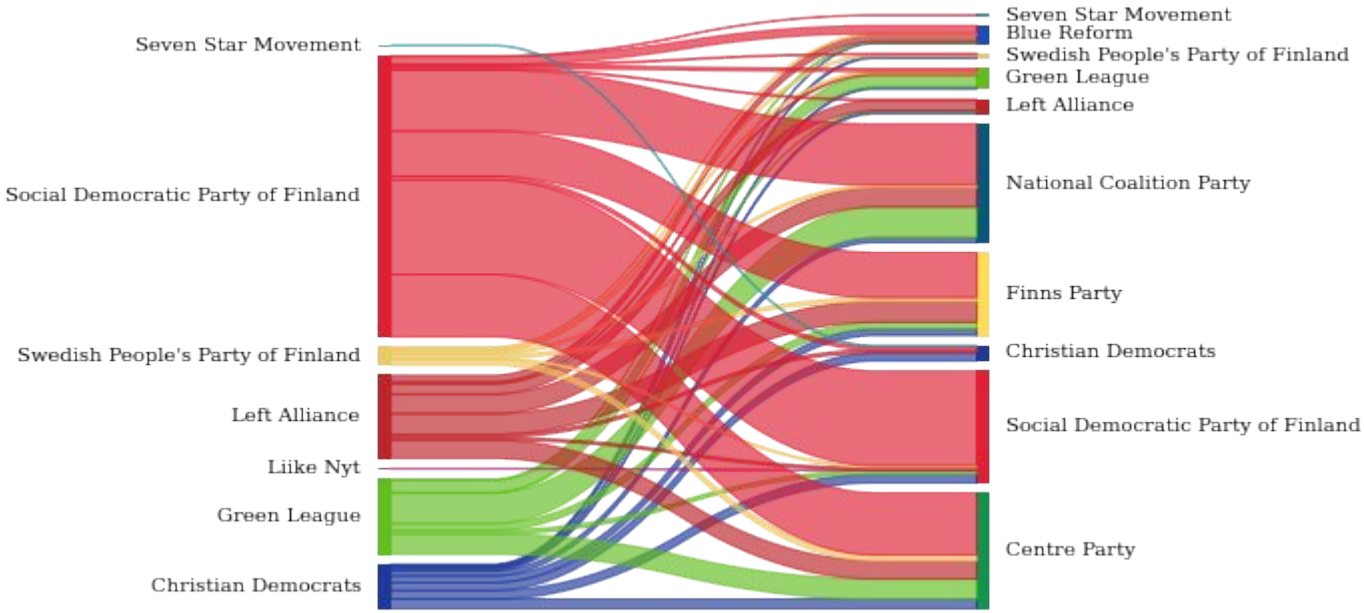
Network Analysis



Network Analysis



Network Analysis



Conclusion

- **Reference network analysis can point out interesting phenomena in parliamentary discussion**
- **Interpreting results requires close reading**
  - ParliamentSampo web portal

Thank you!

More information about the project: https://seco.cs.aalto.fi/projects/semparl/