

Using the LOD service via SPARQL and Notebooks, data analysis, network analysis, and visualizations

Petri Leskinen



Aalto University
School of Science



UNIVERSITY OF HELSINKI



HELDIG

Helsinki Centre for Digital Humanities

HELSINKI INSTITUTE FOR
SOCIAL SCIENCES
AND HUMANITIES

Overview

- SPARQL Protocol and RDF Query Language
 - online in Yasgui editor
 - Queries and Visualizations
- Python
 - Data Analysis
 - Network Analysis

SPARQL in Yasgui

- SPARQL query editor *Yasgui* <https://yasgui.triply.cc/>
 - Perform queries to any publicly open LD service
 - Visualize the results
- *ParliamentSampo* endpoint <https://ldf.fi/semparl/sparql>
 - LOD publication about MPs and Speeches in the Parliament of Finland
 - ... more about ParliamentSampo later today!

SPARQL in Yasgui, simple example

- Example query for Finnish MPs and their parties

<https://api.triplydb.com/s/uncoScdi4>



```
SELECT DISTINCT *  
WHERE {  
  # Instance of the class Person  
  ?person a biocrm:Person ;  
    :party ?party ;  
    skos:prefLabel ?person_label .  
  
  # Full name of the Party  
  ?party skos:prefLabel ?party_label .  
  FILTER (LANG(?party_label) = 'en')  
}
```

SPARQL in Yasgui, simple example

```
13 SELECT DISTINCT *
14 WHERE {
15   # Instance of the class Person
16   ?person a biocrm:Person ;
17   :party ?party ;
18   skos:prefLabel ?person_label .
19
20   # Full name of the Party
21   ?party skos:prefLabel ?party_label .
22   # Show only labels in English
23   FILTER (LANG(?party_label) = 'en')
24 }
```

Press CTRL - <spacebar> to autocomplete

Table Response Gallery Chart Geo Geo-3D Geo events Markup Network Pivot Timeline 3198 results in 0.366 seconds

Simple view ☐ Ellipse ☒ Filter query results Page size: 50

person	party	person_label	party_label
1 <http://ldf.fi/semparl/people/p910494>	group:Q1378704	"Hoikka, Iisakki (1840-1917)"@fi	"Finnish Party"@en
2 <http://ldf.fi/semparl/people/p911770>	group:Q499029	"Wuolijoki, Wäinö (1872-1947)"@fi	"Social Democratic Party of Finland"@en
3 <http://ldf.fi/semparl/people/p911480>	group:Q845537	"Schybergson, Emil (1856-1920)"@fi	"Swedish People's Party of Finland"@en
4 <http://ldf.fi/semparl/people/p911571>	group:Q1713433	"Svinhufvud, P.E. (1861-1944)"@fi	"Young Finnish Party"@en
5 <http://ldf.fi/semparl/people/p911571>	group:Q304191	"Svinhufvud, P.E. (1861-1944)"@fi	"National Coalition Party"@en
6 <http://ldf.fi/semparl/people/p910329>	group:Q1378704	"Danielson-Kalmari, J.R. (1853-1933)"@fi	"Finnish Party"@en
7 <http://ldf.fi/semparl/people/p911490>	group:Q1713433	"Setälä, E.N. (1864-1935)"@fi	"Young Finnish Party"@en

SPARQL in Yasgui

- The parties by the number of members
COUNT, GROUP BY, ORDER BY, DESC

<https://api.triplydb.com/s/aT-p42xqf>

```
SELECT DISTINCT
  ?party ?party_label
  (COUNT(DISTINCT ?person) AS ?number_of_mps)
WHERE {
  ?person a biocrm:Person ;
  :party ?party ;
  skos:prefLabel ?person_label .

  ?party skos:prefLabel ?party_label .
  FILTER (LANG(?party_label) = 'fi')
}
GROUPBY ?party ?party_label
ORDER BY DESC(?number_of_mps)
```

SPARQL in Yasgui, GROUPBY

- <https://api.triplydb.com/s/MPGqgyVIn>

```
12
13 SELECT DISTINCT
14   ?party ?party_label
15   (COUNT(DISTINCT ?person) AS ?number_of_mps)
16 WHERE {
17   ?person a biocrm:Person ;
18   :party ?party ;
19   skos:prefLabel ?person_label .
20
21   ?party skos:prefLabel ?party_label .
22   FILTER (LANG(?party_label) = 'en')
23 }
24 GROUPBY ?party ?party_label
25 ORDER BY DESC(?number_of_mps)
```

Press CTRL - <spacebar> to autocomplete

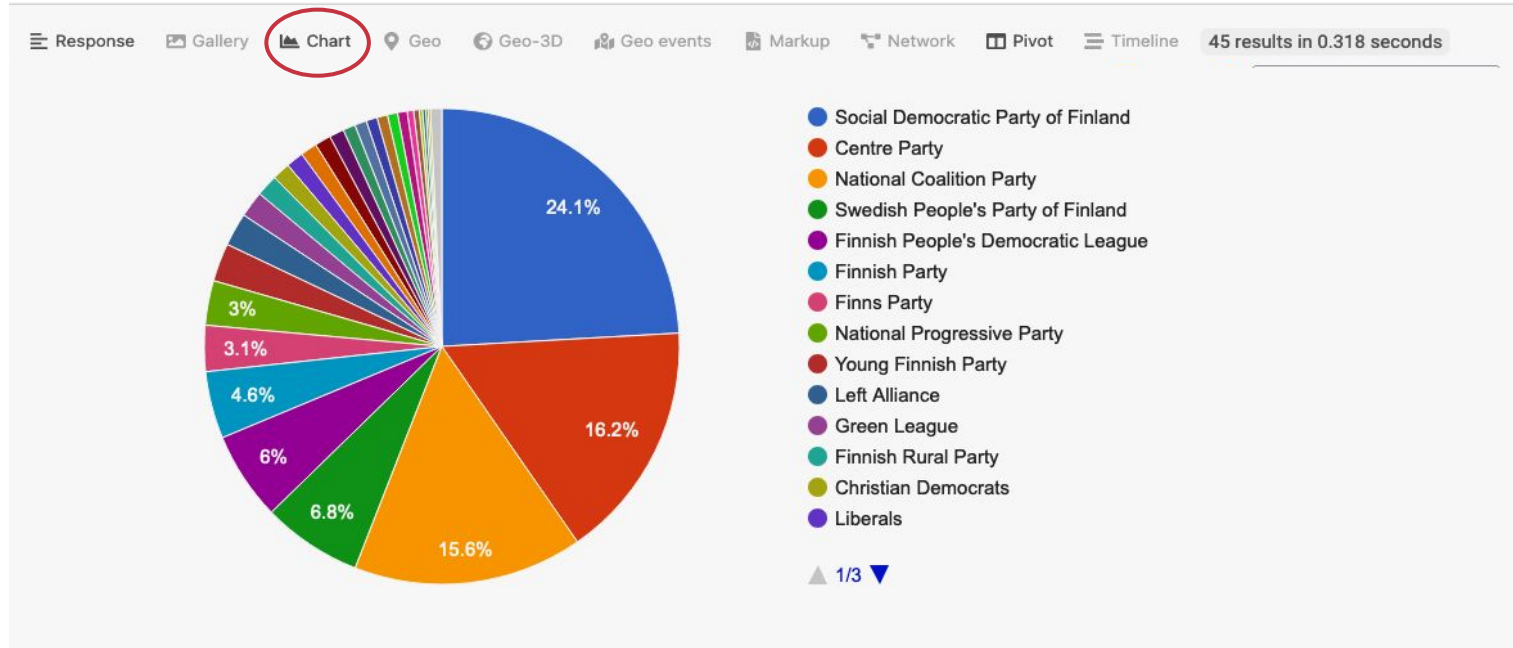
Table Response Gallery Chart Geo Geo-3D Geo events Markup Network Pivot Timeline 45 results in 0.2 seconds

Simple view ☒ Ellipse ☐ Filter query results Page size: 50

party	party_label	number_of_mps
group:Q499029	Social Democratic Party of Finland	771
group:Q506591	Centre Party	519
group:Q304191	National Coalition Party	499
group:Q845537	Swedish People's Party of Finland	217
group:Q585735	Finnish People's Democratic League	192
group:Q1378704	Finnish Party	146
group:Q634277	Finns Party	100
group:Q1191102	National Progressive Party	97
group:Q1713433	Young Finnish Party	83
group:Q385927	Left Alliance	72
group:Q196695	Green League	57

Visualization in Yasgui

- <https://api.triplydb.com/s/31n0EgQVr>



Yasgui, another simple example

- Places of birth: <https://api.triplydb.com/s/Y0inSmRw0>

```
14 SELECT DISTINCT
15   ?fullname
16   ?place_label
17   (CONCAT("Point("+STR(xsd:decimal(?long))+ " "+STR(xsd:decimal(?lat))+")") AS ?coordinate)
18 WHERE {
19   # person name and the event of birth
20   ?mp a biocrm:Person ;
21       skos:prefLabel ?fullname ;
22       crm:P98i_was_born ?birth .
23
24   OPTIONAL {
25     # optional place of birth
26     ?birth crm:P7_took_place_at ?place .
27
28     # the place name and coordinates
29     ?place skos:prefLabel ?place_label ;
30            wgs:lat ?lat ;
31            wgs:long ?long .
32   }
33 }
```

2894 results in 0.525 seconds

fullname	place_label	coordinate
Hoikka, Iisakki (1840-1917)	Kemi	Point(24.58333 65.75)
Wuolijoki, Wäinö (1872-1947)	Hauho	Point(24.56303 61.17255)
Schybergson, Emil (1856-1920)	Turku	Point(22.26869 60.45148)
Svinhufvud, P.E. (1861-1944)	Sääksmäki	Point(24.06575 61.20408)
Danielson-Kalmari, J.R. (1853-1933)	Hauho	Point(24.56303 61.17255)
Setälä, E.N. (1864-1935)	Kokemäki	Point(22.37856 61.30929)

Yasgui, combining earlier examples

- <https://api.triplydb.com/s/WfIG4jti->
- Number of MPs by their party at each geographic location

```
13 SELECT DISTINCT ?place ?place_label ?party ?party_label
14   (COUNT(DISTINCT ?mp) AS ?number_of_MPs)
15   (CONCAT("Point(" + STR(?long) + " " + STR(?lat) + ")") AS ?location)
16   (MIN(?party_color) AS ?locationColor)
17   (CONCAT(?place_label, ' - ', ?party_label) AS ?locationLabel)
18 WHERE {
19   ?mp a biocrm:Person ;
20   crm:P98i_was_born/crm:P7_took_place_at ?place ;
21   :party ?party .
22   ?party skos:prefLabel ?party_label ;
23   :hexcolor ?party_color .
24   FILTER (LANG(?party_label) = 'fi')
25
26   ?place wgs:lat ?lat ;
27           wgs:long ?long ;
28           skos:prefLabel ?place_label .
29 } GROUP BY ?place ?party ?party_label ?place_label
30 ?locationLabel ?long ?lat # HAVING (?number_of_MPs > 1)
31 ORDER BY (?number_of_MPs)
```

Press CTRL - <spacebar> to autocomplete

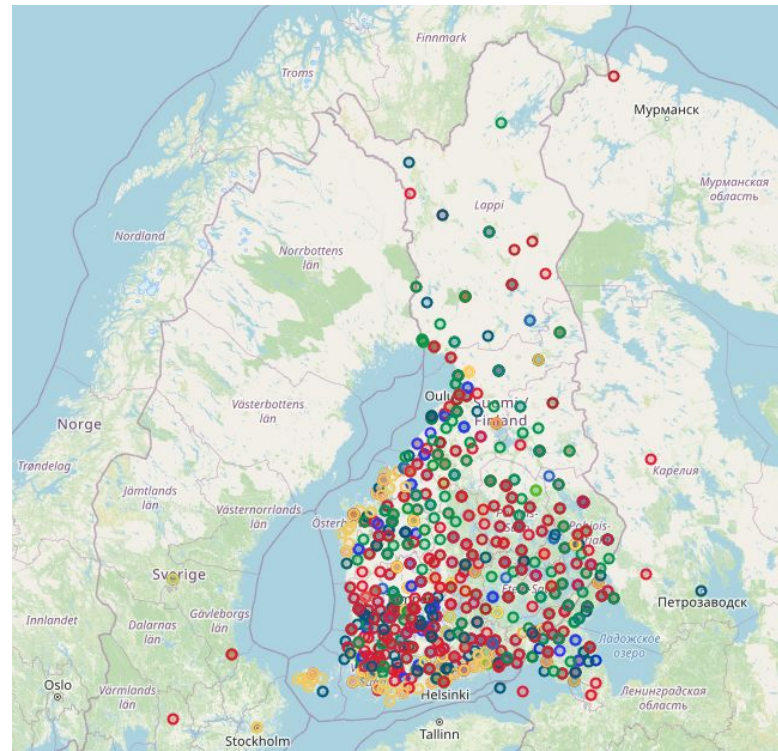
Table Response Gallery Chart **Geo** Geo-3D Geo events Markup Network Pivot Timeline 1279 results in 0.4 seconds

Simple view ☒ Ellipse ☒

place	place_label	party	party_label	number_of_MPs	location	locationColor
http://sws.geonames.org/658225/	Helsinki	group:Q499029	Suomen Sosialidemokraattinen Puolue	55	Point(2.493545e+01 6.016952e+01)	#E11931
http://sws.geonames.org/658225/	Helsinki	group:Q304191	Kansallinen Kokoomus	54	Point(2.493545e+01 6.016952e+01)	#00587C
http://sws.geonames.org/658225/	Helsinki	group:Q845537	Suomen ruotsalainen kansanpuolue	41	Point(2.493545e+01 6.016952e+01)	#EFCB65
http://sws.geonames.org/634963/	Tampere	group:Q499029	Suomen Sosialidemokraattinen Puolue	21	Point(2.378712e+01 6.149911e+01)	#E11931
http://sws.geonames.org/658225/	Helsinki	group:Q196695	Vihreä liitto	21	Point(2.493545e+01 6.016952e+01)	#61BF1A

Places of birth on map

- <https://api.triplydb.com/s/WfIG4jti->
- Place markers colored by the party
 - Swedish People's Party of Finland
 - Centre Party
 - Social Democratic Party of Finland



Data Analysis with Python

- Yasgui allows a lot of simple data analysis and visualization
 - Data tables, Time Series, Pie and Column Charts, Maps
- For more complex and custom analysis
 - Jupyter notebooks or Google Colaboratory
 - Gephi for network analysis

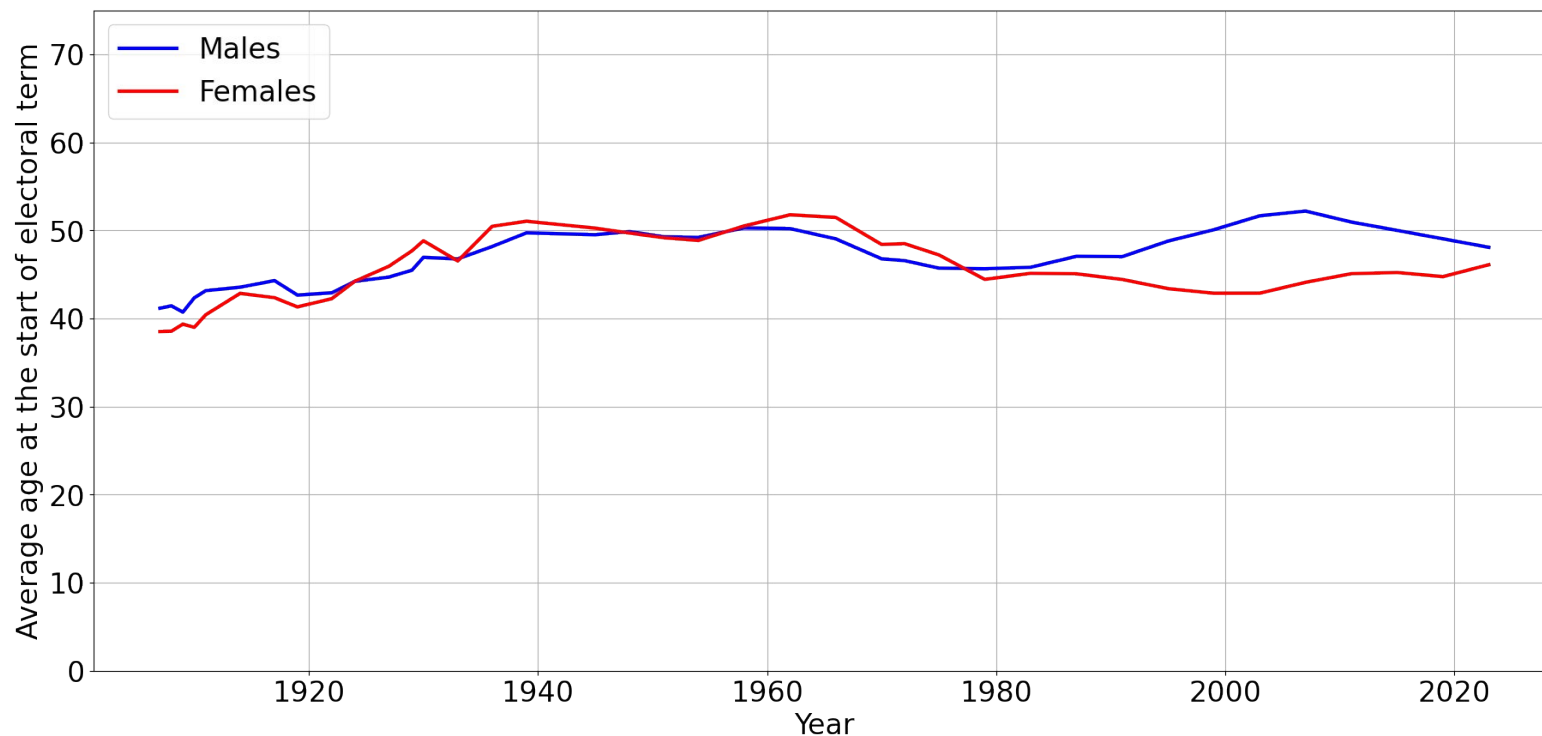
Programming with Python

- Commonly used modules
 - Data acquisition: *RDFLib*, *SPARQLWrapper*
 - Data processing and analysis: *Pandas*, *NumPy*, *SciPy*, *scikit-learn*
 - Visualization: *Matplotlib*, *Seaborn*
 - Networks: *NetworkX*, *Graph-tool*

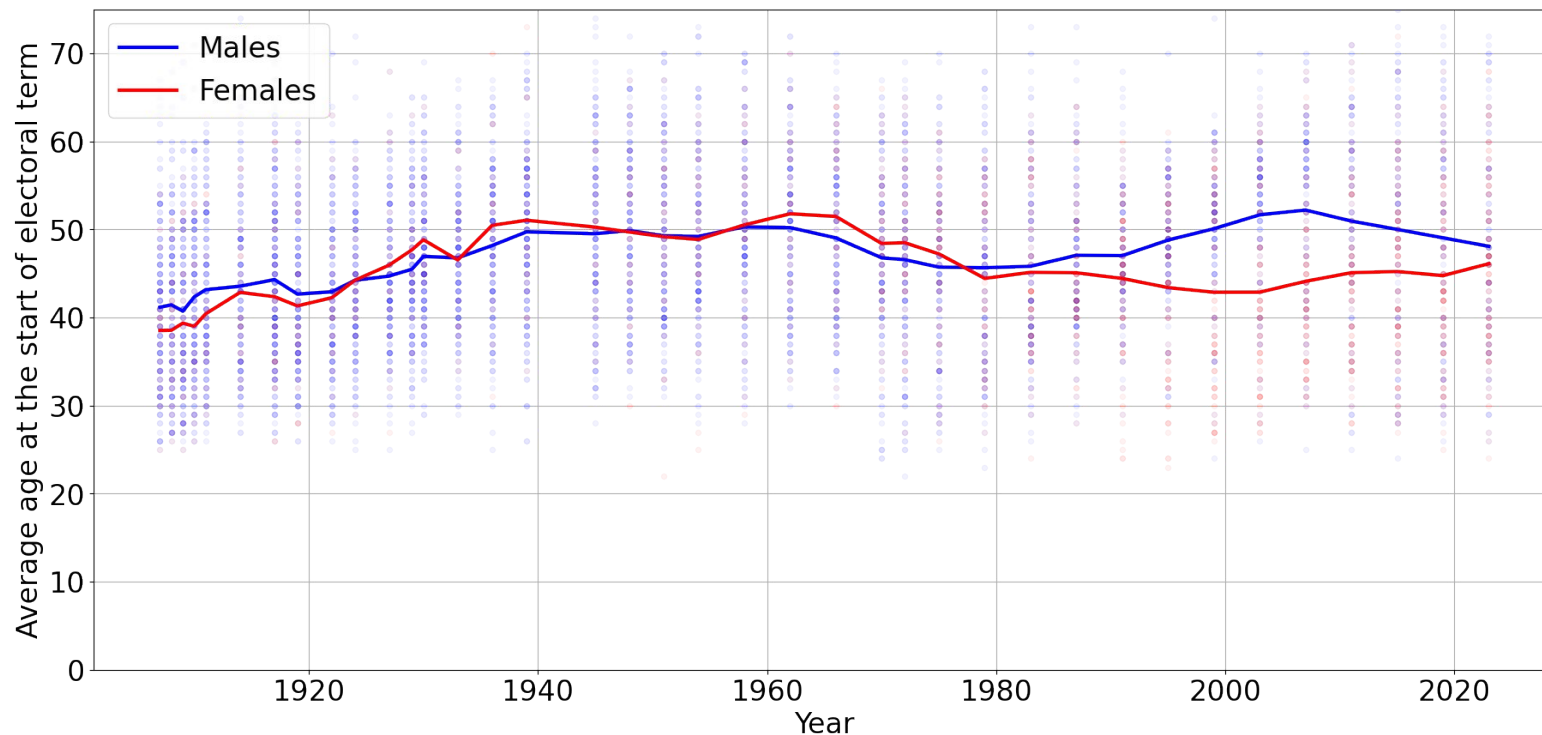
Google Colaboratory Example

- [Google Colaboratory Notebook](#) example using ParliamentSampo data
- Workflow
 - *Install Libraries and import modules*
 - *Query the database*
 - *Process the data*
 - *Output the tables or visualizations*

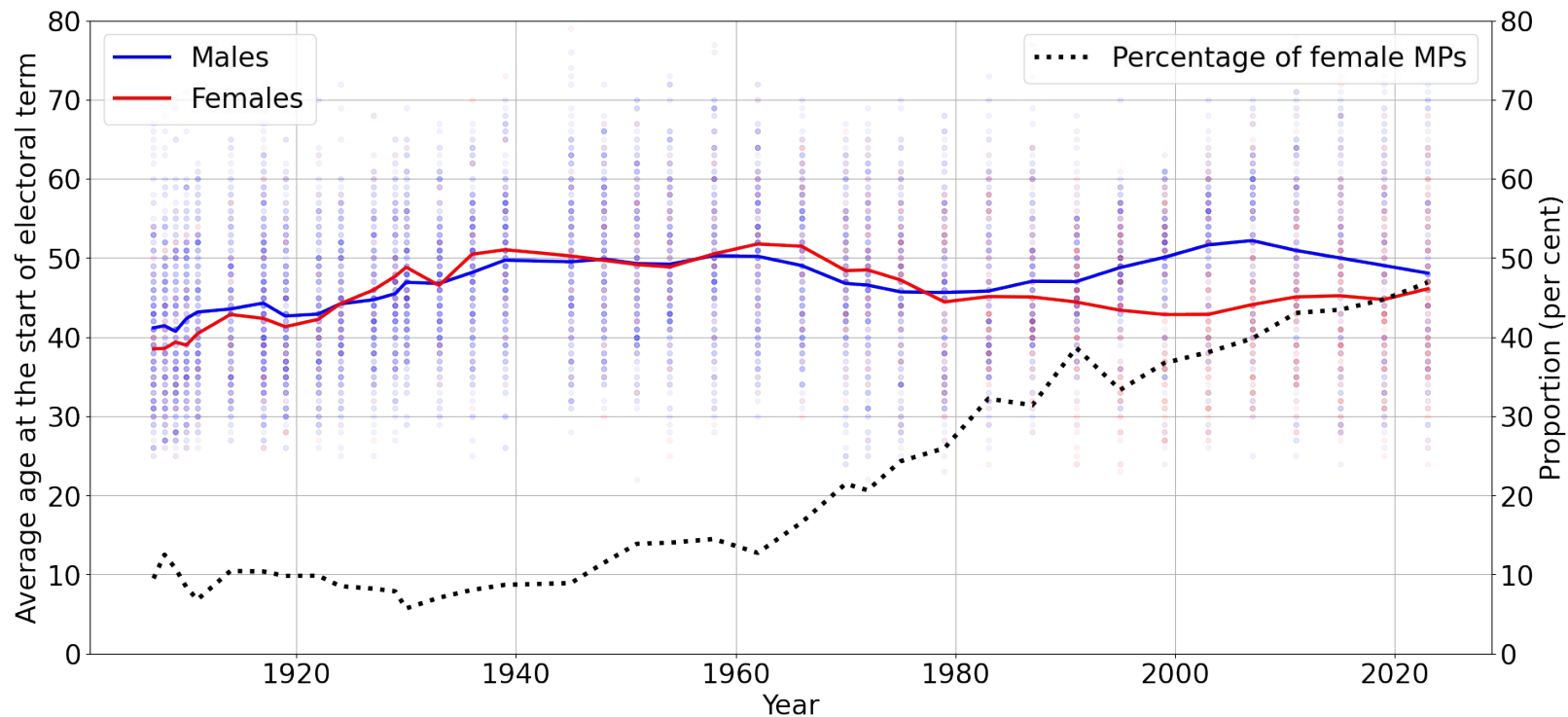
Google Colaboratory Example



Google Colaboratory Example



Google Colaboratory Example



Aalto University
School of Science



UNIVERSITY OF HELSINKI



HELDIG
Helsinki Centre for Digital Humanities

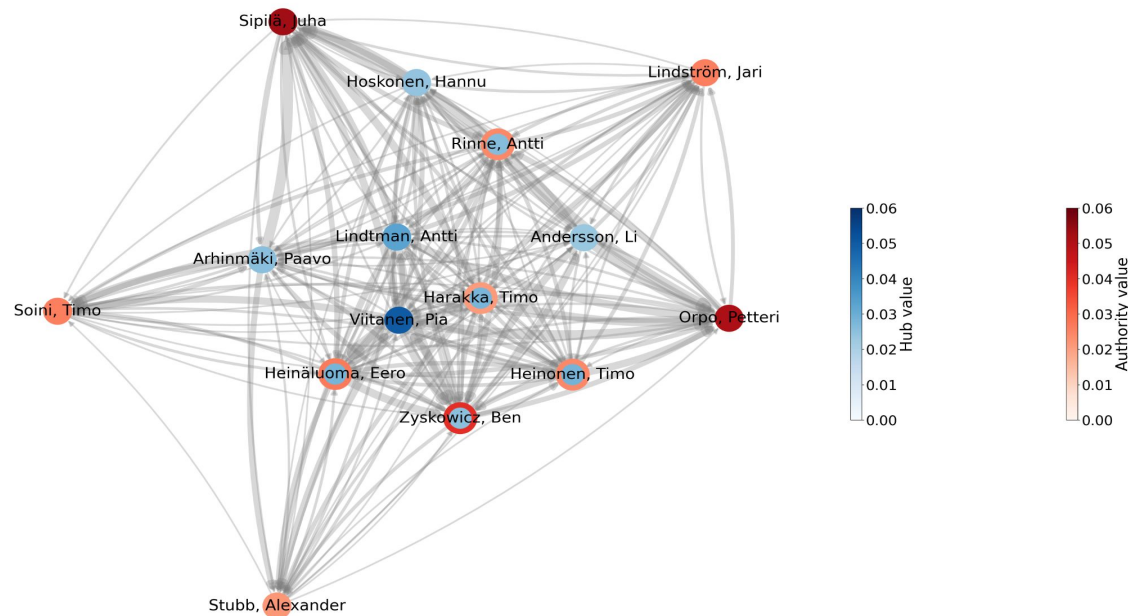
**HELSINKI INSTITUTE FOR
SOCIAL SCIENCES
AND HUMANITIES**

Network Analysis

- Biographical datasets provide networks by multiple criteria
 - Interactions between people
 - *Epistolary data*
 - Reference networks
 - *BiographySampo, ParliamentSampo*
 - Common characteristics
 - *AcademySampo, ParliamentSampo*

Network Analysis

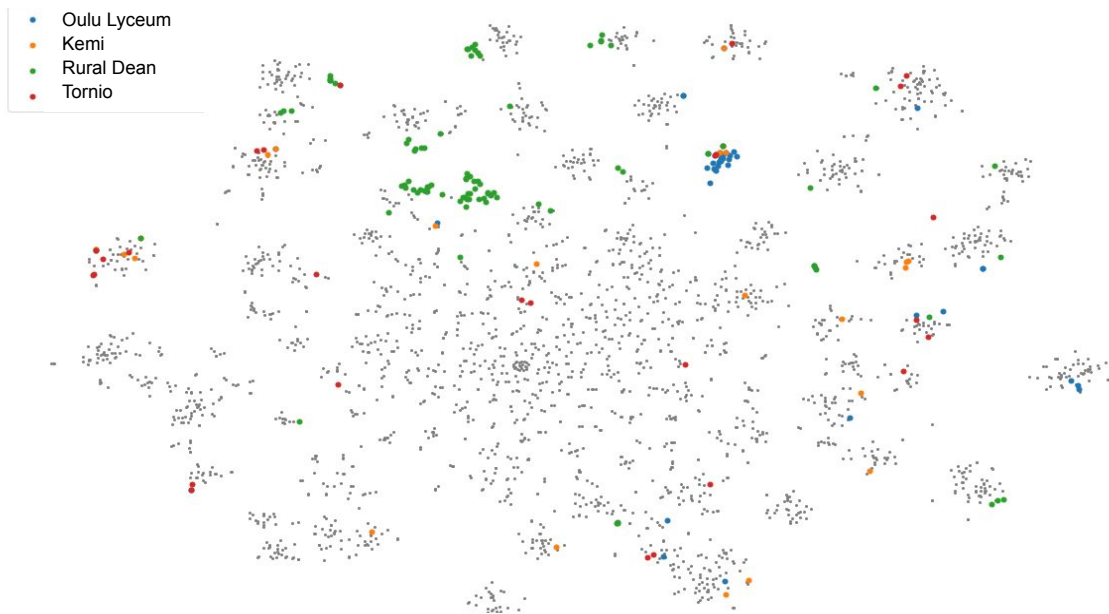
Based on mentioning other people in parliamentary speeches:



Poikkimäki et al.: Applying Network and Bibliometric Analyses to Mentions of Politicians in Plenary Speeches ...

Network and Cluster Analysis

AcademySampo: Embeddings based on lifetime activities, e.g., common vocations, memberships, locations, ...



Recommender System

AcademySampo: Embedding based on lifetime activities, e.g., common vocations, memberships, locations, ...

Recommending actors with similar characteristics in their walks of life on the portal page:

1. Lund, Karl Arne (1878-1919): [Tornion pormestari; Tornio]
2. Grandell, Isak (1691-1732): [Tornion pormestari; Tornio]
3. Ekholm, Nils Gabriel (1758-1809): [Tornion pormestari; Tornio]
4. Lindström, Kaarlo Aleksanteri (1872-1929): [ylempi hallintotutkinto; Yo Oulun lys.; 6.6.1892; Oulun lyseo]
5. Helander (→ Helanen), Olavi Leonard (1869-1931): [kunnallispuormestari; Kemi]
6. Caselius, Ano Eemil (1867-1929): [tuomari; kunnallispuormestari]

Thank You!



Aalto University
School of Science



UNIVERSITY OF HELSINKI



SeCo

HELDIG

Helsinki Centre for Digital Humanities

HELSINKI INSTITUTE FOR
SOCIAL SCIENCES
AND HUMANITIES