

BiographySampo

Biographies on the Semantic Web

Prof. Eero Hyvönen

Aalto University and University of Helsinki
HELDIG – Helsinki Centre for Digital Humanities
Semantic Computing Research Group (SeCo)

Dr. Kirsi Keravuori

Finnish Literature Society

Contents

- **National Biography Collections** of Finnish Literature Society
 - Kirsi Keravuori
- **BiographySampo Portal**
 - AI Reading Biographies for the Semantic Web
 - Paradigm Shift in Publishing Cultural Heritage data
- **BiographySampo Linked Open Data Service**
 - Data Analysis Using Scripting
 - Google Colab, Jupyter, YASGUI etc.





SUOMALAISEN KIRJALLISUUDEN SEURA

Kirsi Keravuori, Finnish Literature Society

NATIONAL BIOGRAPHY COLLECTIONS

Biografiasampo

Suomalaisten elämäkertojen verkosto semantisessa webissä

Biografiasampo mahdollistaa suomalaisten elämäkertojen, henkilöiden ja henkilöryhmien tutkimisen digitaalisten ihmistieteiden menetelmillä. Palvelun ydinaineistonä on Suomalaisen Kirjallisuuden Seuran Kansallisbiografia ja muut tietokannat, yhteensä n. 13 000 elämäkertaa, joita on rikastettu muilla ulkoisilla aineistoilla. Valitse alta sovellusnäkymä aineistoihin. Lisätietoa [ohjeessa](#).

Valitse sovellusnäkymä aineistoihin



Biographical databases

Etsi henkilöä nimellä **Hae**

**Biografiakeskus –
Kansallisbiografia
ja paljon muuta**

VERKKOJULKAIKSET: **KANSALLISBIOGRAFIA** **TALOUSELÄMÄN VAIKUTTAJAT** **HENKILÖKUVAT** **KENRAALIT JA AMIRAALIT**
TURUN HIIPPAKUNNAN PAIMENMUISTO **SKS:N JÄSENET** **SUOMEN PAPISTO 1800–1920**

Kokoelma	määrä
Kansallisbiografia	6500
Kenraalit ja amiraalit 1809-1917	481
Suomen papisto 1800-1920	1234
Talouselämän vaikuttajat	2235
Turun hiippakunnan paimenmuisto 1554-1721	2716
Yhteensä elämäkertoja	13166

Viikon henkilö

Algoth Untola

Algoth Untola on ollut pitkään ristiriitainen hahmo suomalaisessa kirjallisuudessa. Toisaalta hänet on tuomittu vallankumouksellisen ideologiansa takia, toisaalta hänen Maiju Lassilan nimellä kirjoittamansa humoristiset kertomukset ja kansannäytelmät sekä niiden todenoisiset ihmiset ja puheet ovat vuosikymmenien ajan huvittaneet koko Suomea. Algoth Untolan syntymästä on 28.11.2018 kulunut 150 vuotta.

[Lue lisää](#)

Ajankohtaista

Kansallisbiografiaan lisää elokuvan ja teatterin tekijöitä

Kansallisbiografia-verkkojulkaisua on täydennetty uusilla elokuva- ja teatterialan henkilöiden elämäkerroilla. Kokosalaisen on lisätty muun muassa ruotsalaisen teatterin...

[Lue lisää](#)

Biografiataso avattu

Suomalaisen elämäkertojen ja niitä rikastavan datan verkopalvelu Biografiataso on avattu osoitteessa <http://biografiataso.fi>. Digitalista humanistista tutkimusta...

[Lue lisää](#)

Jo yli 2000 Lönnrotin kirjettä avoimesti verkossa!

Suurin osa Lönnrotin lähettämistä

[Henkilön tiedot](#)[Kartta ja aikajana](#)[Verkosto](#)[Viitaukset](#)[Kielianalyysi](#)

Blanka Namurilainen

Tutki henkilöön liittyvää verkostoa

Sivulla näytetään havainnekuva henkilöön liittyvästä verkosta. Linkitys on luotu biografiakuvausten viittausten perusteella.

Verkoston solmuja voidaan siirtää hiirellä. Solmuja klikkaamalla pääset tarkastelemaan verkostoa valitun henkilön ympärillä. Alapuoleilla valkoilla voidaan säätää verkoston laajuutta ja ulkoasua.

Näytetään 194 linkkiä ja 50 henkilöä

Näytä verkostossa korkeintaan 50 henkilöä.

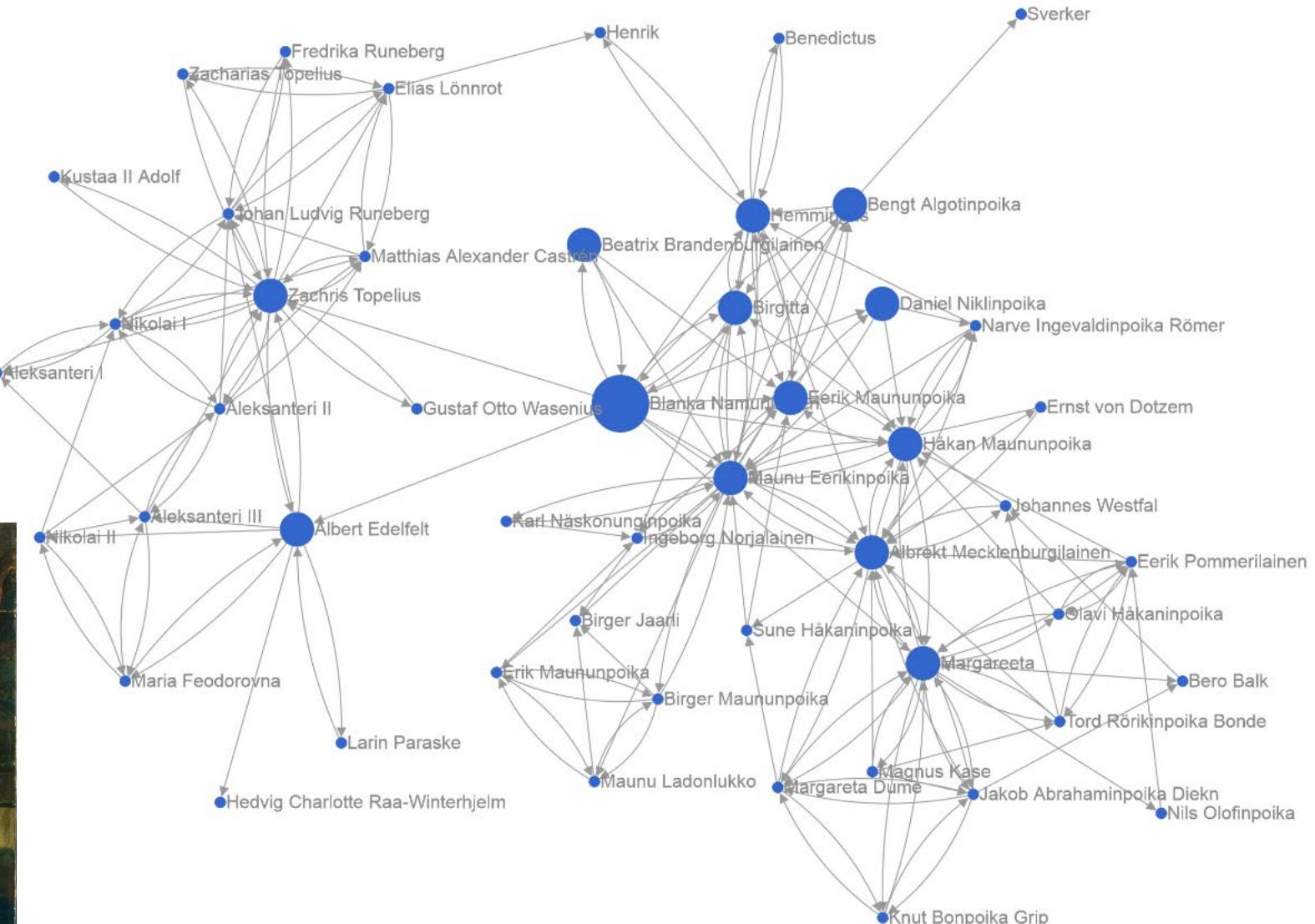
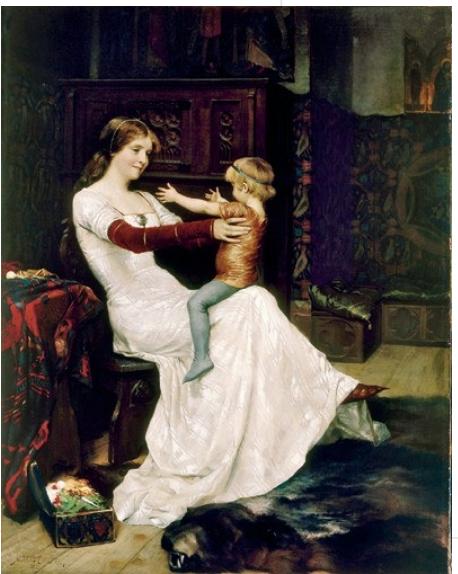
Verkostossa näytetään

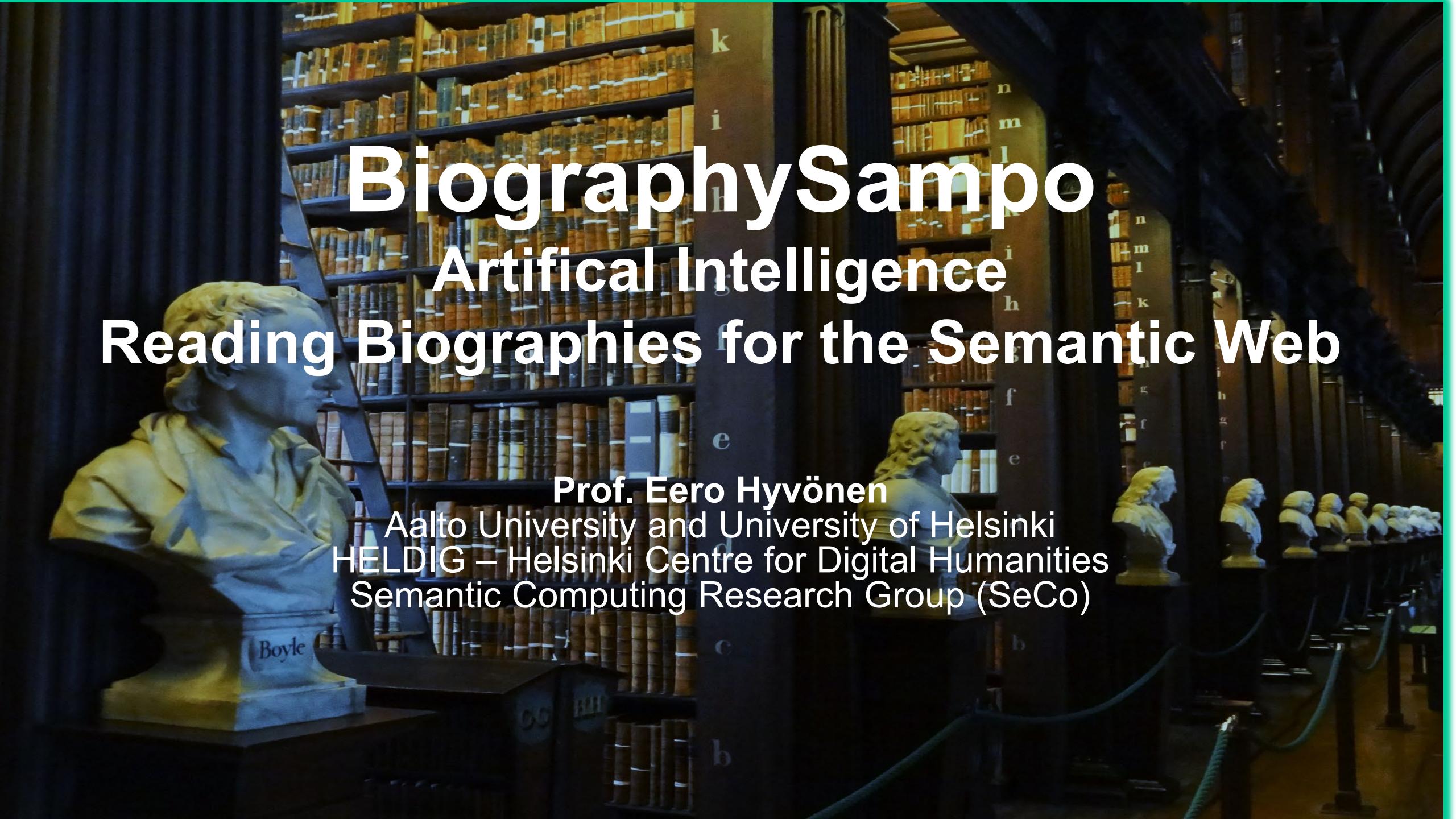
Aineistoon käsin kirjatut linkit

Automaattisesti tunnistetut linkit

Solmujen koko: Etäisyys

Solmujen väri: Etäisyys





BiographySampo

Artifical Intelligence

Reading Biographies for the Semantic Web

Prof. Eero Hyvönen

Aalto University and University of Helsinki
HELDIG – Helsinki Centre for Digital Humanities
Semantic Computing Research Group (SeCo)

BiographySampo: Artifical Intelligence Reading Biographies for the Semantic Web



vimeo Manage videos Watch Features Stock NEW Upgrade Search videos, people, and more Upload

The video player displays a scene from a library. In the foreground, several busts of historical figures are visible on a shelf. The background shows tall bookshelves filled with books. Overlaid on the image is the title 'BiographySampo' in large, bold, white letters, followed by the subtitle 'Artifical Intelligence Reading Biographies for the Semantic Web' in smaller white text. A play button and a timestamp '05:10' are at the bottom left. A progress bar is at the bottom right.

BiographySampo - Finnish Biographies on the Semantic Web NOT YET RATED

2 months ago | More

 SeCo Research Group

Settings Review tools Send file

This column is only visible to you

See all video stats [View Stats](#)

Related Videos Autoplay next video

 BiographySampo - ...
SeCo Research Group

<https://vimeo.com/328419960>

BiographySampo Demonstrates a Paradigm Shift: 4 Generations of Publishing Data for Humanities

1. Printed Texts
2. Online Systems for Searching and Exploring
3. Publishing Content as Linked Data with Tools for DH
4. Automatic Knowledge Discovery and Artificial Intelligence

1. GENERATION SYSTEMS

PRINTED TEXTS



National Biography of Finland

ca. 13100 bios



SUOMALAISEN KIRJALLISUUDEN SEURA

2. GENERATION SYSTEMS

TEXTS ONLINE



Haku

Tällä hakulomakkeella voit hakea kaikkia Kansallisbiografian elämäkertoja. Artikkeleita on lähes 6 500, ja kokoelmaa päivitetään ja täydennetään jatkuvasti.

Kaikkien artikkeleiden tiivistelmää ja satoja [vapaasti luettavia artikkeleita](#) voit lukea ilmaiseksi. Vapaasti luettavat artikkelit näkyvät hakutuloksissa vihreällä väriillä. Muiden artikkelien lukeminen edellyttää, että organisaatioi on hankkinut käyttölisenssin. Yksityishenkilöt voivat lukea artikkeleita joko hankkimalla määräaikaisen käyttäjätunnuksen tai ostamalla yksittäisiä artikkeleita APE-kukkarolla.

[Hanki käyttäjätunnus](#)

Selaa henkilötä sukunimen alkukirjaimen mukaan:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Ä	Å	O
Suku- ja/tai etunimi																												
Ammatti																												
Toimiala																								Mikä tahansa				
Kokoelma																								<input checked="" type="checkbox"/> Kansallisbiografia				
																								<input checked="" type="checkbox"/> Kansallisbiografia II				
Aikakausi																								<input checked="" type="checkbox"/> Ruotsin vallan aika				
																								<input checked="" type="checkbox"/> Suuriruhtinaskunnan aika				
																								<input checked="" type="checkbox"/> Tasavallan aika				
Syntymävuosi																								<input type="button" value="▼"/>	-	<input type="button" value="▼"/>		
Kuolinvuosi																								<input type="button" value="▼"/>	-	<input type="button" value="▼"/>		
Syntymäkunta																												
Syntymämaakunta																								Mikä tahansa				
-maa																								Mikä tahansa				
Kuolinkunta																												
Kuolinmaakunta																								Mikä tahansa				
-maa																								Mikä tahansa				
Sukupuoli																								Mikä tahansa				
Kirjoittaja																								Mikä tahansa				
Hakusana(t)																												

<https://kansallisbiografia.fi/>

Biographies online
for close reading

Biografiakeskus

Kansallisbiografia

[Etusivu](#) > [Kansallisbiografia](#) > Saarinen, Eliel (1873 - 1950)

[Vapaasti luettava artikkeli](#)

Saarinen, Eliel (1873 - 1950)

arkkitehti, professori



Aarne Pietinen. Uuden
Suomen kuva-arkisto.

Eliel Saarinen saavutti 1800- ja 1900-luvun vaihteessa kansainvälistä mainetta Suomeen suunnittelemaan jugendtyylisillä rakennuksilla sekä näyttelyarkkitehtina. Hänen Helsinkiä koskevissa kaupunkisuunnitelmissaan hahmoteltiin pääkaupungille kasvusuuntia, jotka ovat toteutuneet vasta viime aikoina. Saarinen siirtyi 1920-luvulla Yhdysvaltoihin, missä hän suunnitteli muun muassa pilvenpiirtäjiä ja toimi arkkitehtuurin opettajana.

Eliel Saarisen arkkitehdin ura kesti vuodesta 1896 vuoteen 1950. Se jakaantui melkein tasaisesti kahteen yli neljännesvuosisadan pitaiseen jaksoon: Suomen-aikaan ja Amerikan-aikaan. Suomessa Saarinen saavutti melkein heti arkkitehtuurin uudistajan maineen ja nousi 1900-

Hae kaikista julkaisuista

Etsi henkilöä nimellä



Kansallisbiografia

- [Haku](#)
- [Esittely](#)
- [Artikkelen rakenne ja lähteet](#)
- [Ohjeita käyttäjälle](#)
- [Ohjeita kirjoittajille](#)
- [Tekijänoikeudet](#)

- [Palaute](#)
- [Vapaasti luettavat](#)
- [Kirjaudu sisään](#)

<https://kansallisbiografia.fi/>

3. GENERATION SYSTEMS

TEXTS AS (LINKED) DATA

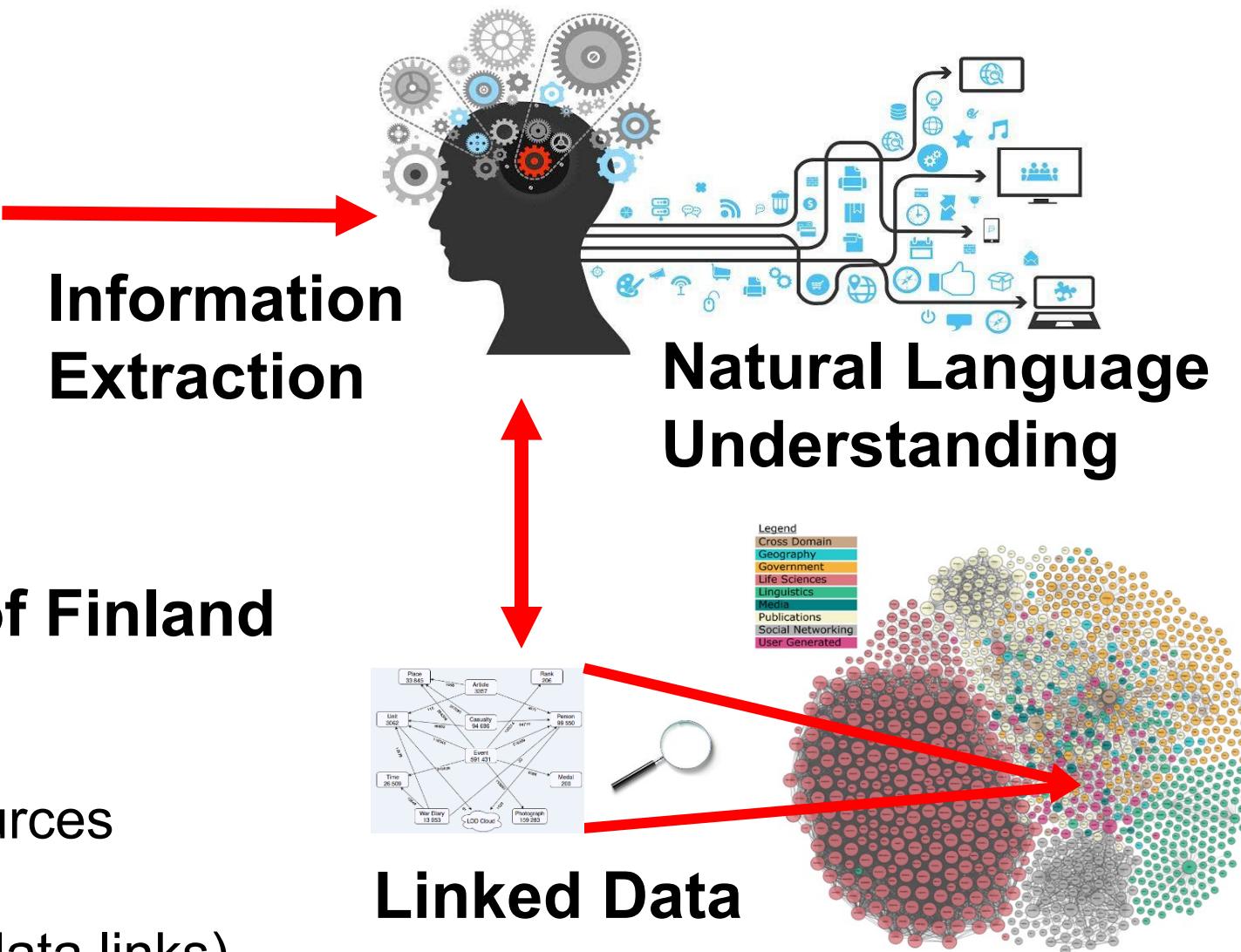
TOOLS FOR DATA ANALYSIS

Artificial Intelligence Reading Biographies for the Semantic Web!



National Biographies of Finland

- 13 144 biographies
- Written by 977 authors
- Interlinked with 16 data sources
- 164 000 events
- Over 120 000 000 triples (data links)



Data-analytic tools: 19th century generals & admirals vs. clergy

Explore biographies maps

Map combined with the persons **native** and **died** places with arrows. If the place is the same, it is shown as a circle. On the lower side filters may be limited to a desired group of persons.

By clicking on the arrows and circles on a map, a window opens, showing a list of persons who have events associated with location.

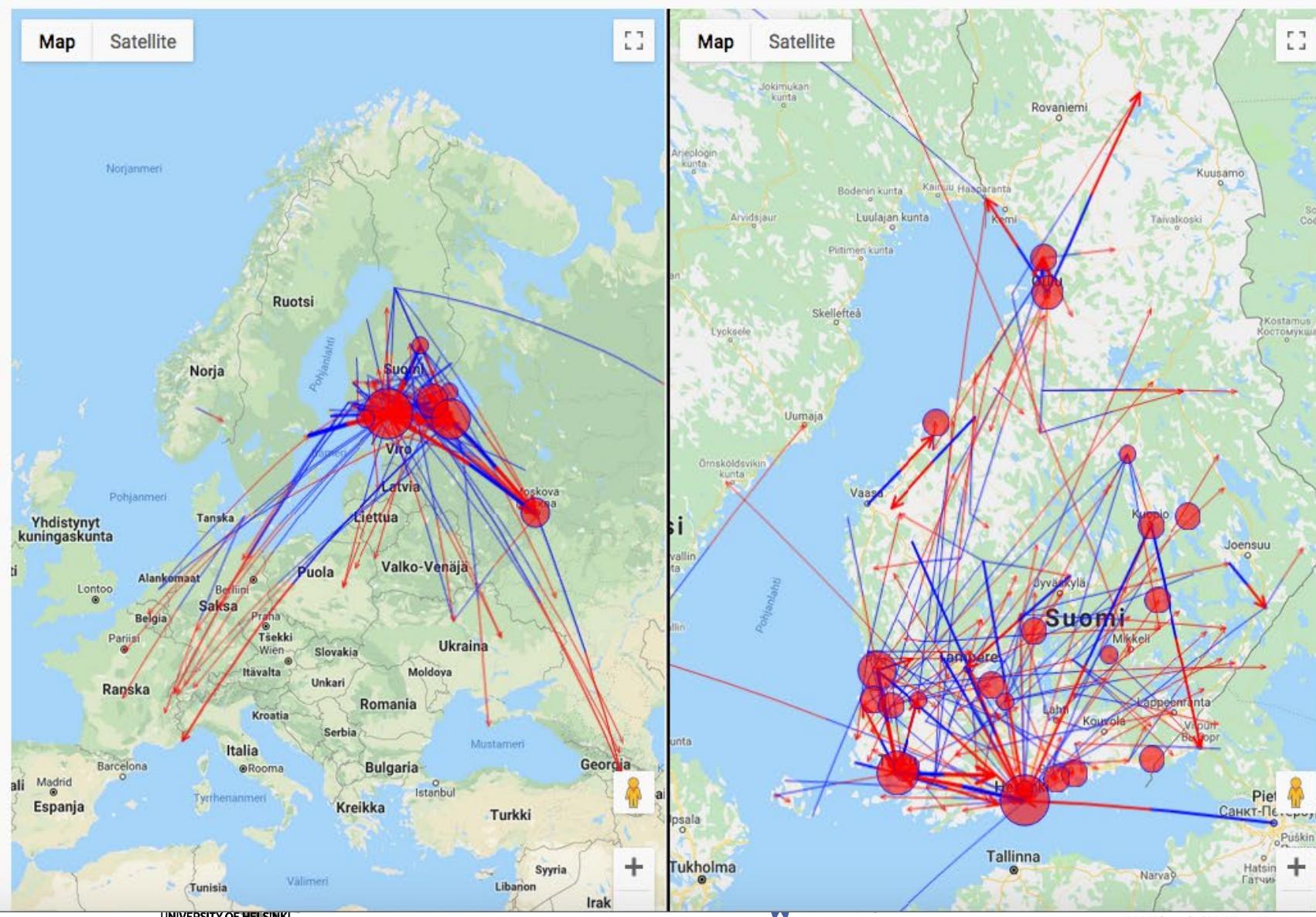
[View the map up to 200 results.](#)

The search returned 199 results place.

[Remove all selections](#)

Limit the date of birth of persons
200 2020

Database
- No selection - (481)
The generals and admirals



[View the map up to 200 results.](#)
Search results yielded 198 place.

[Remove all selections](#)

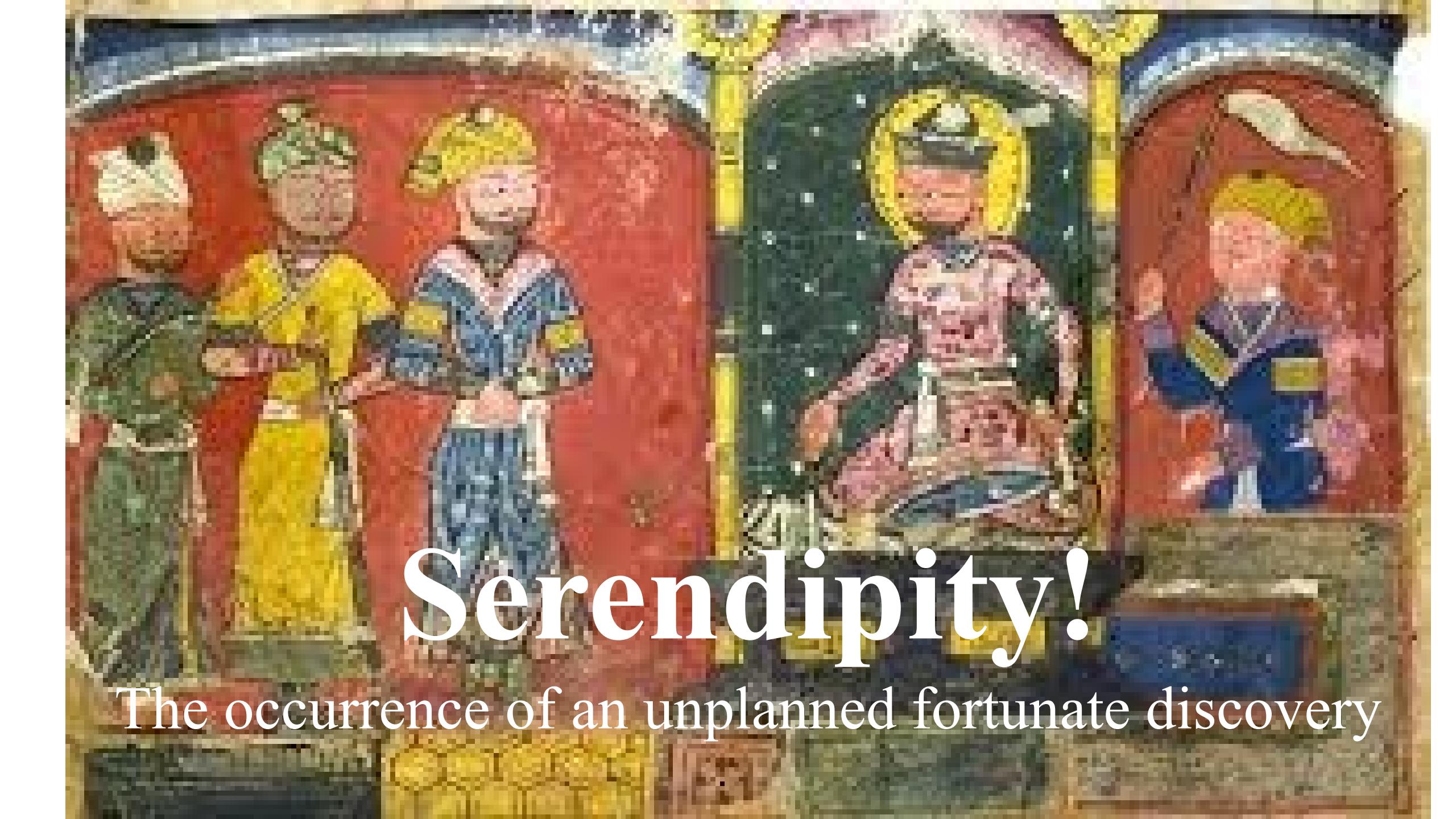
Limit the date of birth of persons
200 2020

Database
- No selection - (1234)
Finnish Clergy 1800-1920

linked databases
 Wikipedia (123)
 Wiki data (120)
 Fenicia - The National Finland (114)
 Military Sampo (1)
 Norssit (18)
 Host Sampo (12)
 Biografiskt Lexikon for Finland (4)
 ULAN (0)
 VIAF (35)
 Geni.com (749)
 Home (0)
 The Parliament (19)

4. GENERATION SYSTEMS

AUTOMATIC KNOWLEDGE DISCOVERY & AI



Serendipity!

The occurrence of an unplanned fortunate discovery

Publication System as an Intelligent Agent

- Passive Tools -> Active intelligent agents
 - Agents **find** "interesting" research questions in the data
 - Agents **solve** research questions
 - Agents can **explain** their solutions

Semantic Web 0(0) 1
IOS Press

Using the Semantic Web in Digital Humanities: Shift from Data Publishing to Data-analysis and Serendipitous Knowledge Discovery

Eero Hyvönen
*University of Helsinki, Helsinki Centre for Digital Humanities (HELDIG), Finland and
Aalto University, Department of Computer Science, Finland*
Semantics Computing Research Group (SeCo) (<http://seco.cs.aalto.fi>)
E-mail: eero.hyvonen@aalto.fi

Abstract. This paper discusses a shift of focus in research on Cultural Heritage semantic portals, based on Linked Data, and envisions and proposes new directions of research. Three generations of portals are identified: Ten years ago the research focus in semantic portal development was on data harmonization, aggregation, search, and browsing ("first generation systems"). At the moment, the rise of Digital Humanities research has started to shift the focus to providing the user with integrated tools for solving research problems in interactive ways ("second generation systems"). This paper envisions and argues that the next step ahead to "third generation systems" is based on Artificial Intelligence: future portals not only provide tools for the human to solve problems but are used for finding research problems in the first place; for addressing them, and even for solving them automatically under the constraints set by the human researcher. Such systems should preferably be able to explain their reasoning, which is an important aspect in the source critical humanities research tradition. The second and third generation scenarios call now a challenge for both computer scientists and humanities researchers.

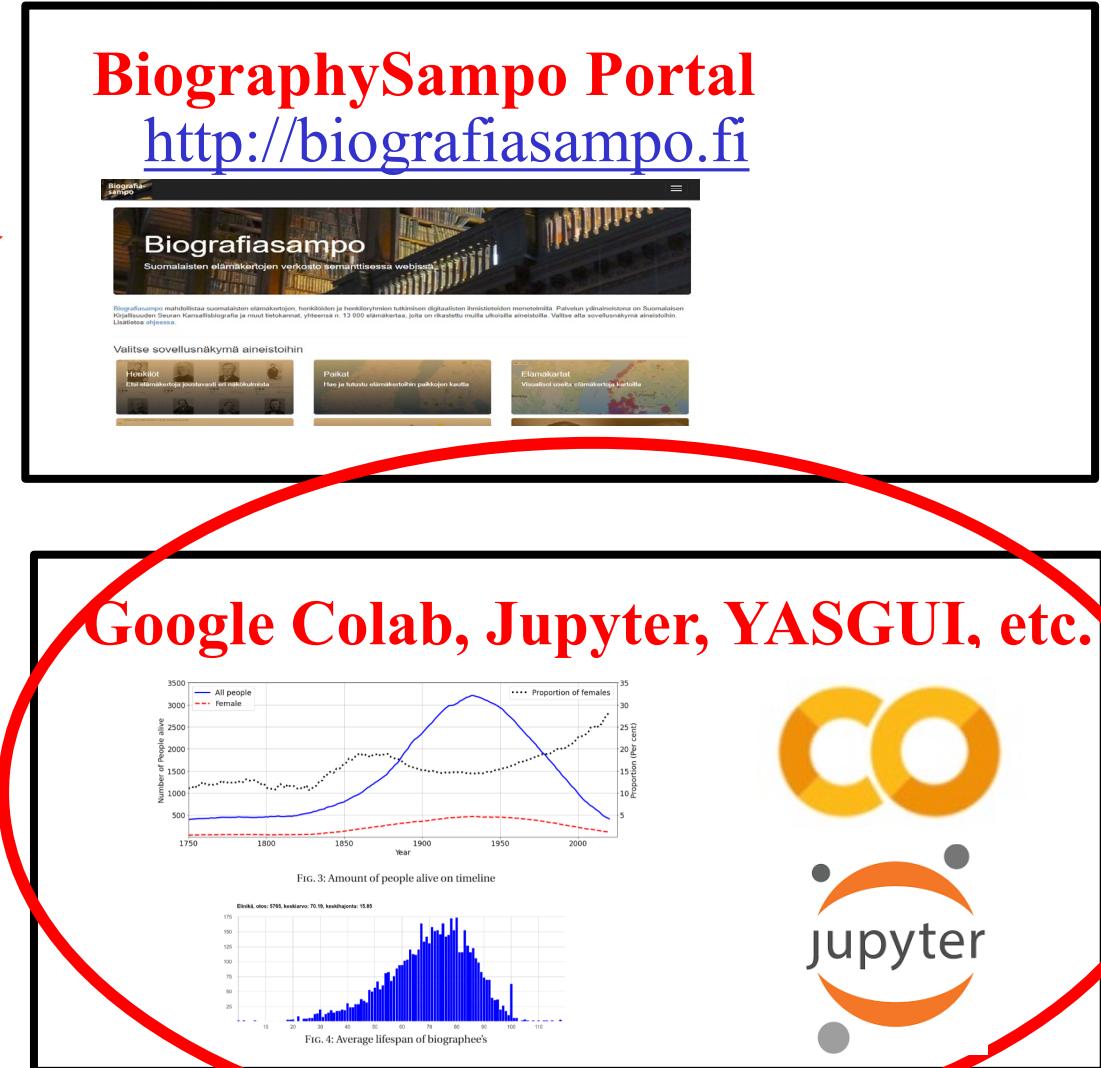
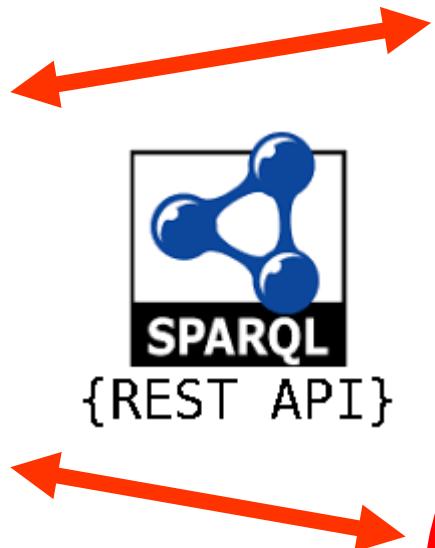
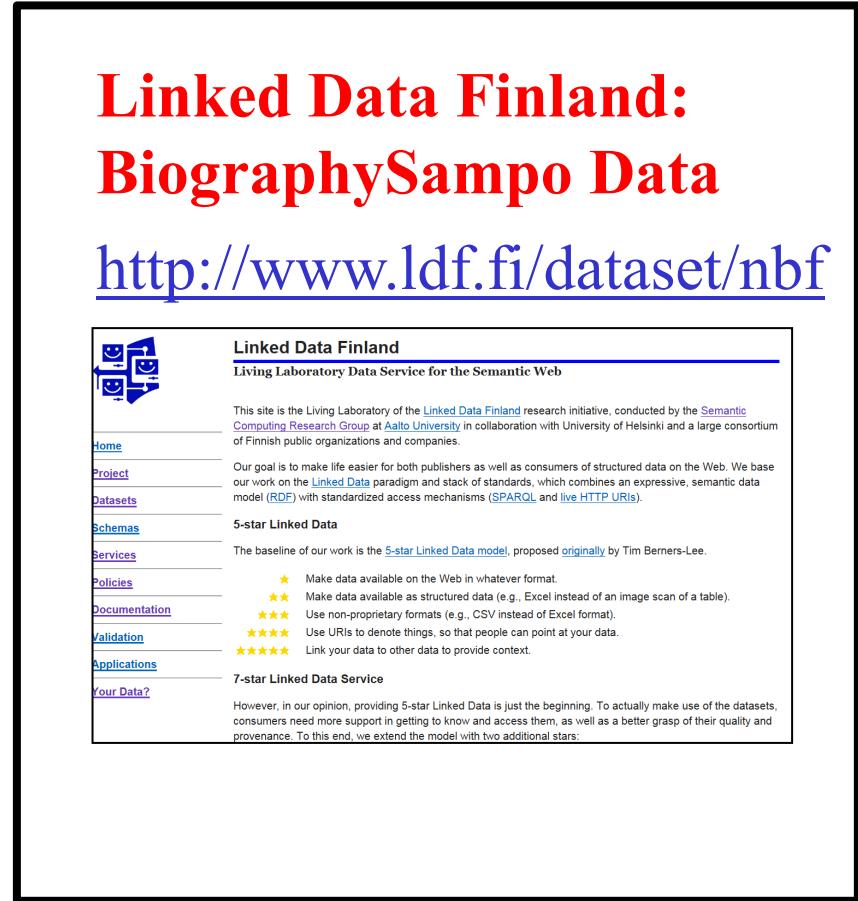
Keywords: Digital Humanities, Linked Data, Semantic portals, Data analysis, Knowledge discovery

1. Introduction

and forms a substantial part of DBpedia³ and Wikidata⁴.

The availability of Big Data has boosted the rapidly emerging new research area of Digital Humanities

Two Faces of BiographySampo: for 1) Machines and 2) Human Users



Analyzing Biography Collections Historiographically as Linked Data: Case National Biography of Finland

Minna Tamper^{1,2[0000-0002-3301-1705]}, Petri Leskinen^{1[0000-0003-2327-6942]}, and
Eero Hyvönen^{1,2[0000-0003-1695-5840]}

¹ Semantic Computing Research Group (SeCo), Aalto University, Finland

² HELDIG – Helsinki Centre for Digital Humanities, University of Helsinki, Finland

<http://seco.cs.aalto.fi>, <http://heldig.fi>
firstname.lastname@aalto.fi





Table of contents X

- BiographySampo visualization queries
 - Support functions
 - category and occupation lookups
- Quering person data
 - average number of categories and occupations for person
 - Cumulative sum of people alive
 - Number of people on a timeline
 - Common occupations
 - table by genders
 - Timeline of occupations
 - 80/20 rule
 - Parents' occupations
 - Years of events
 - Person–Parent categories
 - Biographee–Parent occupations
 - Link differences in age
 - Events
 - most common countries/continents
 - Category correlations
 - self categories
 - correlations of rewarded
 - + Section

+ Code + Text

```
ax.legend(['All people', 'Female'], loc = 'upper right')

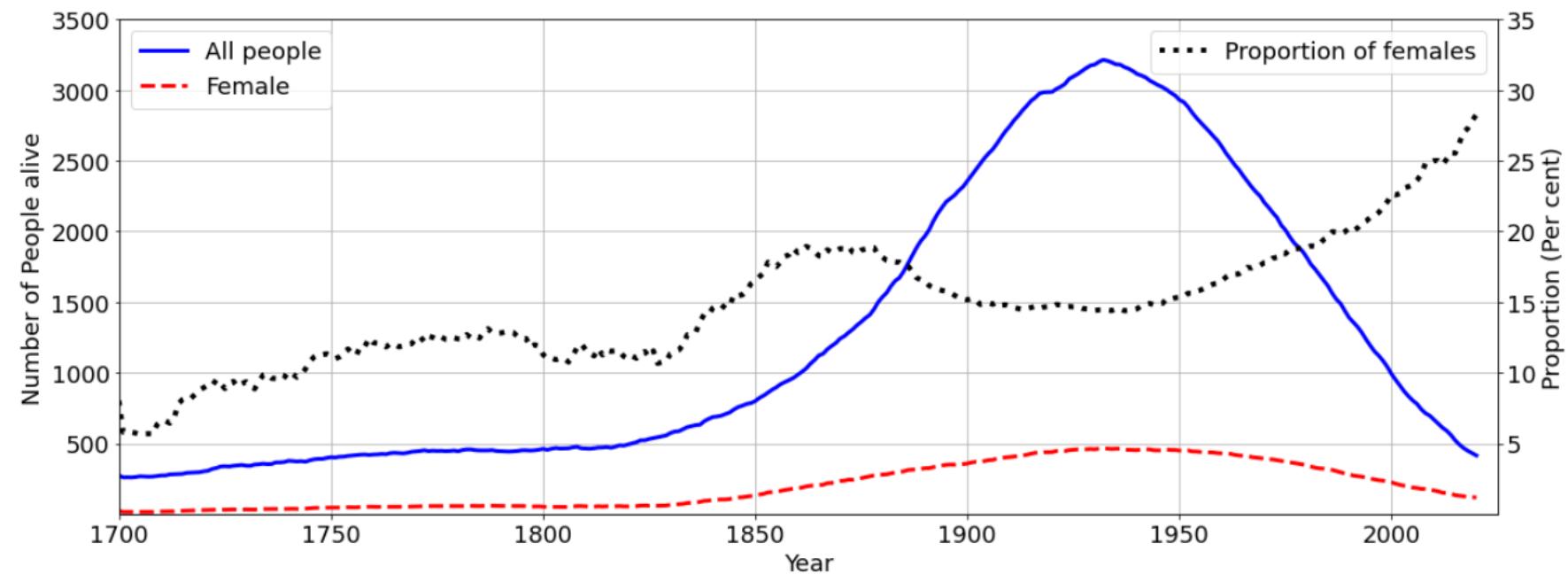
ax.grid()
plt.xlabel('Year')
plt.ylabel('Number of People alive')
plt.tick_params(labelright=True)

ax2 = ax.twinx()
h3=ax2.plot(list(XYR.keys()), list(XYR.values()), 'k:', linewidth=4)
plt.ylim(0.01, 35)

ax2.set_ylabel('Proportion (Per cent)', color='k')
ax2.tick_params('y', colors='k')

ax2.legend(['Proportion of females'], loc = 'upper right')
# ax.legend((h1,h2,h3),('All people', 'Female', 'Proportion'))

plt.savefig("Female_proportion.png", dpi=None, transparent=False, bbox_inches='tight')
plt.show()
```



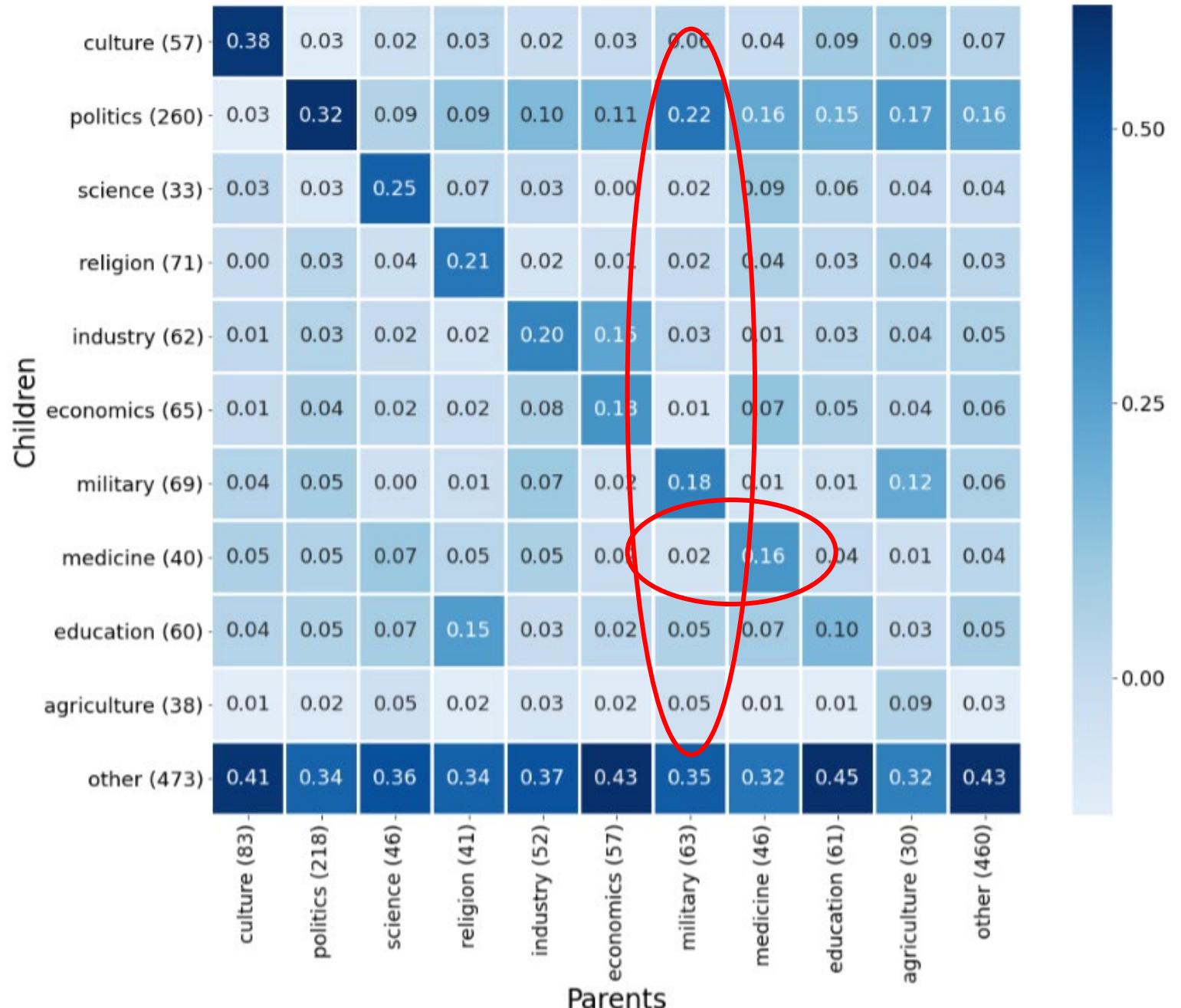


FIG. 10: Correlations between the vocational groups of parents and children

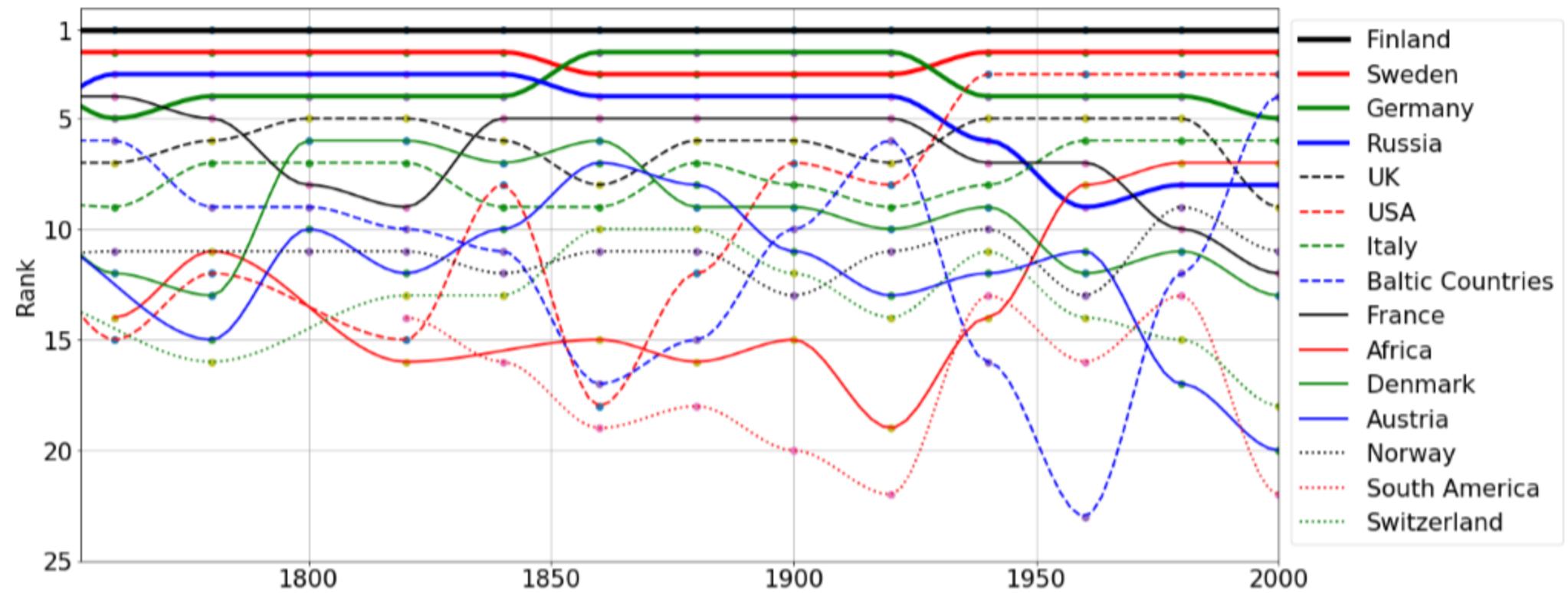


FIG. 14: Top 15 countries on a timeline

More Info – Questions?

Linked Open Infra for Digital Humanities in Finland: LODI4DH

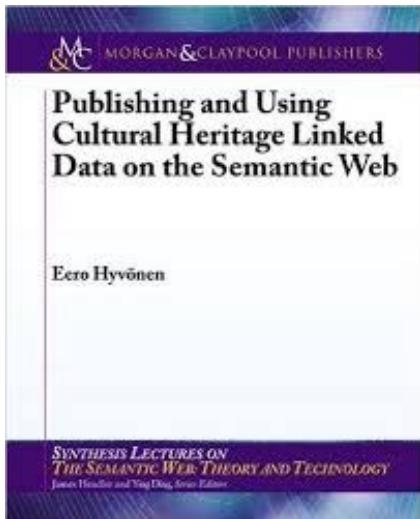
<https://seco.cs.aalto.fi/projects/lodi4dh/>

Semantic Web & Linked Data

<http://www.w3.org/standards/semanticweb/>

Sampo Model & Applications

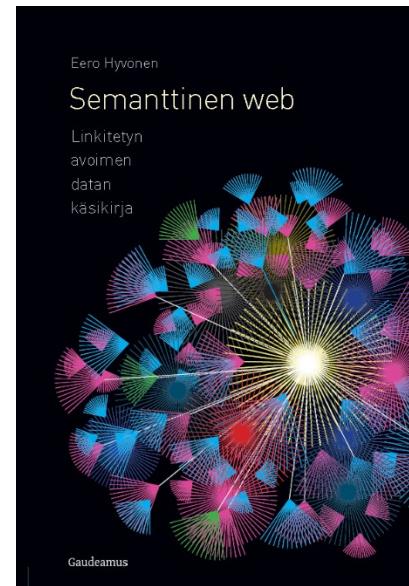
<http://seco.cs.aalto.fi/publications>



In English

2012

<https://www.amazon.com/Publishing-Cultural-Heritage-Synthesis-Technology/dp/1608459977>



In Finnish

2018

<https://www.gaudeamus.fi/semanttininen-web/>