



# Four Generations of Publishing and Using Texts in Digital Humanities: Forging Sampo Portals in the Digital Age

Eero Hyvönen

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

# Contents

- Vision: Four generations of publishing and using texts
- Example: BiographySampo and "Sampo" series of portals

# Four generations of publishing and using texts



# 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  
Semantic 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 systems set new challenges for both computer scientists and humanities researchers.

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

## 1. Introduction

Cultural Heritage (CH) has become a most active

and forms a substantial part of DBpedia<sup>3</sup> and Wikidata<sup>4</sup>.

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

# **Four Generations of Publishing and Using Texts in DH**

- 1. Printed Texts**
- 2. Online Systems for Publishing Texts**
- 3. Publishing and Using Texts as Linked Data**
- 4. Automatic Knowledge Discovery and Artificial Intelligence**

# BiographySampo: AI Reading Biographies for the Semantic Web



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

BiographySampo  
Artifical Intelligence Reading Biographies  
for the Semantic Web

05:10

BiographySampo - Finnish Biographies on the Semantic Web NOT YET RATED

2 months ago | More

SeCo Research Group

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>

# **1. GENERATION SYSTEMS**

**PRINTED TEXTS**



# National Biography of Finland

ca. 13100 bios



## SUOMALAISEN KIRJALLISUUDEN SEURA



# Akatemiasampo

# Historialliset akateemiset henkilöt Suomessa 1640-1899

In English

## Projektin tausta

Helsingin yliopiston keskusarkiston alkuperäisaineistoista digitoidut Ylioppilasmatrikkeli sekä lähdeaineistoihin perustavat matrikkeliit 1640-1899 sisältävät yksityiskohtaisia tietoja n. 28 000 historiallisesta suomalaisesta, käytönnössä kaikista tuon aikakauden akateemisen koulutuksen saaneista henkilöistä maassamme. Pitkälti käsitönä tehty n. 10 henkilötövuoden suuryö matrikkeliiden digitoinnissa, translitteroinnissa ja täydentämisessä kattaa keskeisen osan Suomen ja suomalaisen yliopistolaitoksen historiaa. Suomalaisen Kirjallisuuden Seuran, Suomen Historiallisen Seuran ja muiden tieteellisten seurojen yhteistyönä on puolestaan luotu Kansallisbiografia ja muita biografiakokoelmia, yhteensä 13 144 pienoiselämäkertaa 977 elämäkerturin voimin. Aineisto on ollut saatavilla suurteosarjana ja sitä täydentävinä tietokantoina ja biografiat on nyttemmin julkaistu myös verkkopalveluna. Syksyllä 2018 aineistosta valmistui yhteistyössä Aalto-yliopiston ja Helsingin yliopiston digitaalisten ihmistieteiden keskuksen HELDIG kanssa kieli-, web- ja tekoliitetyknologiaa hyödyntävä [Biografiataso-palvelu](#). Sen kansainvälisenä innovaationa on tarjota lukijalle sisällöllisesti rikastettujen elämäkertojen ohella ylälineitä data-analyyttistä elämäkerrallista ja prosopografista tutkimusta varten.

Aukeama alkuperäisestä Ylioppilasmatrikkeliista 1640-1852

## Akatemiasampo-hankkeen tavoitteet

Akatemiasampo-hankkeen tavoitteena on Helsingin yliopiston ylioppilasmatrikkeliin 1640-1899 rikastaminen niihin liittyvien muiden aineistojen avulla ja julkaiseminen älykkääänä Biografiatasossa avoimena verkkopalveluna. Palvelu on tarkoitettu tutkijoiden, yliopistolaisten ja laajemman yleisön käytettäväksi ja hyödynnettäväksi tutkimushankkeissa ja uusien sovellusten kehittämisen yhteydessä. Akademiasampo tuo kuvaan omasta historiastaan ylpeästä, teknisesti edistyneestä suomalaisesta yliopistolaitoksesta digitaalisten ihmistieteiden keinoin. Järjestelmän kehittämiseen sisältyy omaa metodologista mielenkiintoa: siiä tutkitaan eri lähteistä saatavien

<https://seco.cs.aalto.fi/projects/yo-matrikkelit/>

## **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 värisillä. 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

Q

## Kansallisbiografia

[Haku](#)

[Esittely](#)

[Artikkelen rakenne ja lähteet](#)

[Ohjeita käyttäjälle](#)

[Ohjeita kirjoittajille](#)

[Tekijänoikeudet](#)

[Palaute](#)

[Vapaasti luettavat](#)

[Kirjaudu sisään](#)

<https://kansallisbiografia.fi/>



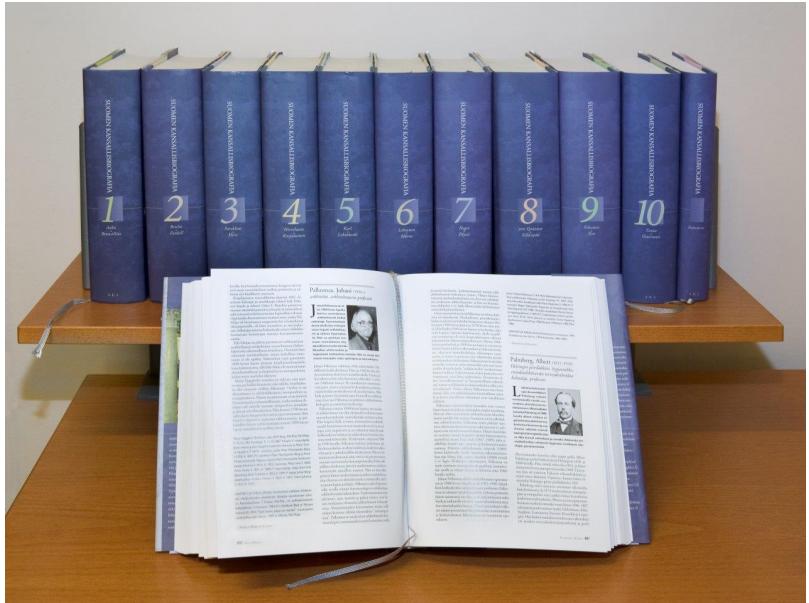
**DAXIA**

The image shows a large, ornate library with floor-to-ceiling bookshelves filled with books. A prominent red 'X' is drawn across the center of the image, covering the word 'DAXIA'. The letters are large and white, standing out against the dark background of the library.

## **3. GENERATION SYSTEMS**

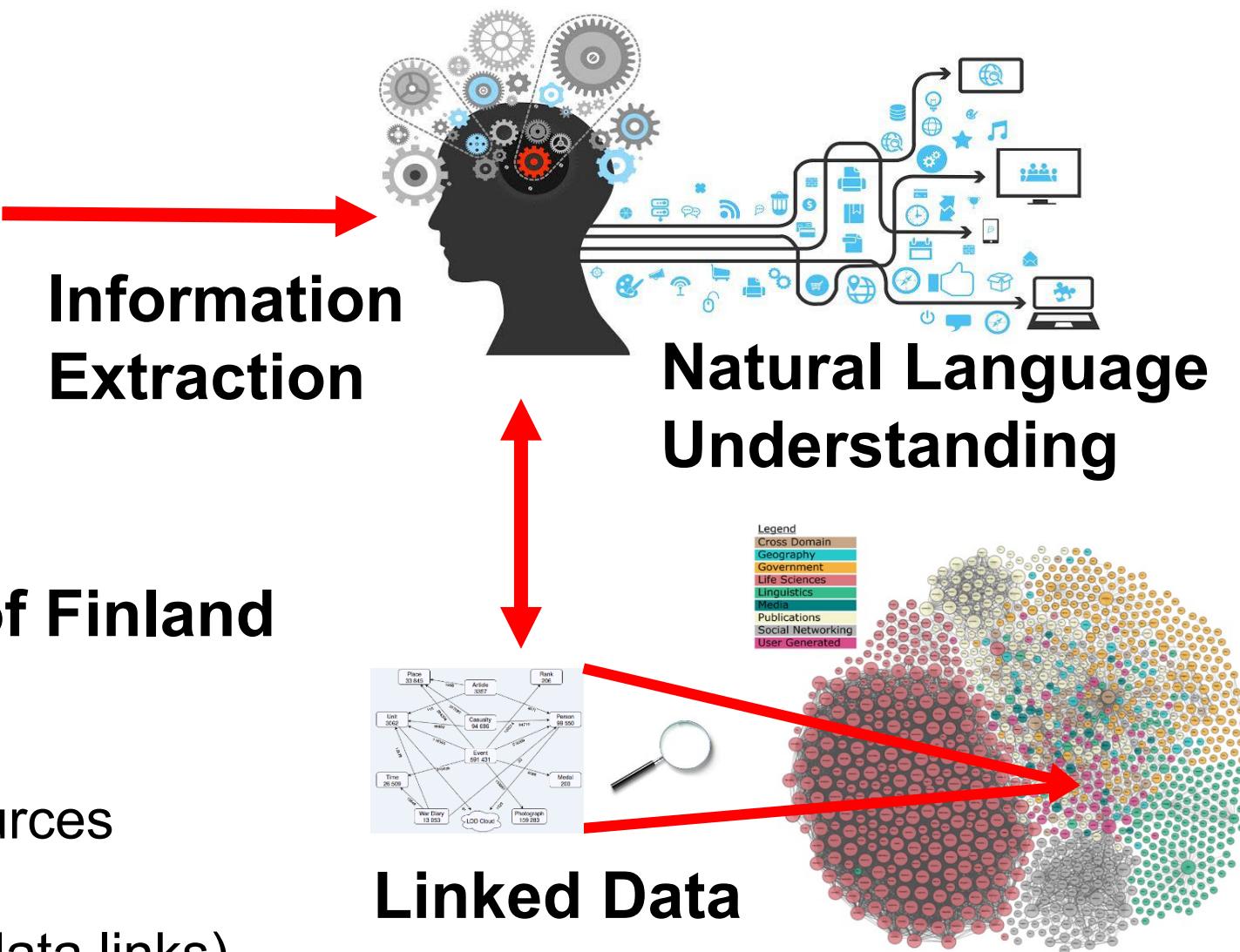
**TEXTS AS (LINKED) DATA**

# 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.

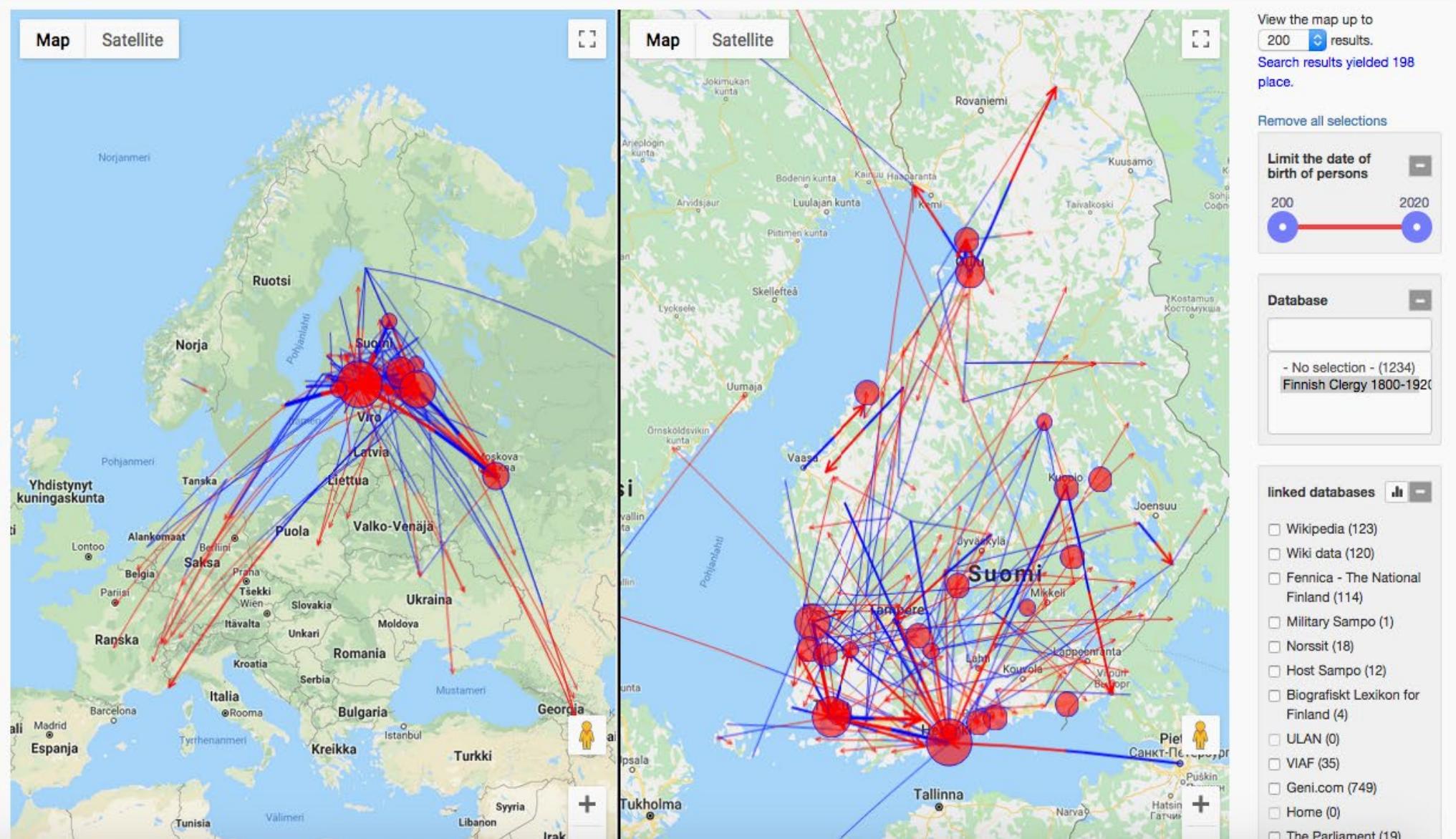
[View the map up to 200 results.](#)

The search returned 199 results place.

[Remove all selections](#)

**Limit the date of birth of persons**  
200

**Database**  
- No selection - (481)  
[The generals and admirals](#)



## **4. GENERATION SYSTEMS**

**AUTOMATIC KNOWLEDGE DISCOVERY & AI**

# Publication System as an Intelligent Agent

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

# Example: Relational Search in BiographySampo

## “How are Finnish Artists Related to Italy?”

The screenshot shows the BiographySampo website's search interface. The top navigation bar includes links for Henkilöt, Paikat, Elämäkartat, Tilastot, Verkostot, Yhteyshaku, and Kielianalyysi. The search bar contains the query "Taidemaalareiden yhteyksiä Firenzeen". Below the search bar, there are three highlighted search results: "Kalle Päätalon yhteyksiä Pohjois-Pohjanmaalle", "J. L. Runebergin yhteyksiä paikoihin", and "Taidemaalareiden yhteyksiä Firenzeen". The main content area displays a table of search results with columns for Henkilö, Paikka, Lähde, and Yhteyden tyyppi. The results list various connections between Finnish artists and Firenze, such as Elin Danielson-Gambogi's connection to the Accademia di Belle Arti and her work on the 'Porta della Mandorla'.

Poista kaikki valinnat

**Henkilö**

- No Selection (24)
  - Danielson-Gambogi, Elin (1861 - 1919) (3)
  - Enroth, Erik (1917 - 1975) (1)
  - Favén, Mauri (1920 - 2006) (1)
  - Gebhard, Albert (1869 - 1937) (1)
  - Harsti, Reino (1900 - 1979) (3)
  - Järnefelt, Eero (1863 - 1937) (1)
  - Liipola, Yrjö (1881 - 1971) (1)
  - Schjerbeck, Helene (1862 - 1946) (1)
  - Sievänen, Jaakko (1932 - 2013) (1)

**Henkilön arvo, ammatti tai toiminta**

Persons = ?

Occupation = Artist

Place = Italy

— Connection type =?

Answers in  
Finnish

| Yhteyden kuvaus  | Henkilö                               | Paikka  | Lähde  | Yhteyden tyyppi                   |
|--|---------------------------------------|---------|--|-----------------------------------|
| Elin Danielson-Gambogi on vastaanottanut kunnianosoitukseen joka liittyy paikkaan Firenze: 'Firenzen kaupungin taidepalkinto 1899'.            | Danielson-Gambogi, Elin (1861 - 1919) | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Kunnianosoitus liittyy paikkaan   |
| Elin Danielson-Gambogi on luonut vuonna 1902 taideokseen 'Firenzen seudulta' kuvaa paikkaa Firenze.  | Danielson-Gambogi, Elin (1861 - 1919) | Firenze | Taulun tiedot Kansallisgallerian tietokannassa | Maalaus kuvaa paikkaa             |
| Elin Danielson-Gambogi on luonut vuonna 1902 taideokseen 'Firenzen seudulta' kuvaa paikkaa Firenze.  | Danielson-Gambogi, Elin (1861 - 1919) | Firenze | Taulun tiedot Kansallisgallerian tietokannassa | Maalaus kuvaa paikkaa             |
| Henkilön Erik Enroth uraan tai opiskeluun liittyy paikassa Firenze tapahtuma: Accademia di Belle Arti, Firenze 1957                            | Enroth, Erik (1917 - 1975)            | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Ura tai opiskelu liittyy paikkaan |
| Henkilön Mauri Favén uraan tai opiskeluun liittyy paikassa Firenze tapahtuma: Accademia di Belle Arti, Firenze, Italia 1958                    | Favén, Mauri (1920 - 2006)            | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Ura tai opiskelu liittyy paikkaan |
| Henkilön Albert Gebhard uraan tai opiskeluun liittyy paikassa Firenze tapahtuma: Scuola Libera, Firenze  | Gebhard, Albert (1869 - 1937)         | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Ura tai opiskelu liittyy paikkaan |
| taideokseen 'Ikkunastani Firenzessä', joka kuvaa paikkaa Firenze.  | Harsti, Reino (1900 - 1979)           | Firenze | Taulun tiedot Kansallisgallerian tietokannassa | Maalaus kuvaa paikkaa             |
| taideokseen 'Sypressejä (Firenze)', joka kuvaa paikkaa Firenze.  | Harsti, Reino (1900 - 1979)           | Firenze | Taulun tiedot Kansallisgallerian tietokannassa | Maalaus kuvaa paikkaa             |
| taideokseen 'Portaali Firenzessä', joka kuvaa paikkaa Firenze.   | Harsti, Reino (1900 - 1979)           | Firenze | Taulun tiedot Kansallisgallerian tietokannassa | Maalaus kuvaa paikkaa             |
| taideokseen 'Taitelijan pojien muotokuva (Heikki Järnefelt)', joka kuvaa jun liittyy paikassa Firenze tapahtuma: opintoja Firenzessä, Roomassa | Järnefelt, Eero (1863 - 1937)         | Firenze | Taulun tiedot Kansallisgallerian tietokannassa | Maalaus kuvaa paikkaa             |
| opiskeluun liittyy paikassa Firenze tapahtuma: matka Wieniin,  | Liipola, Yrjö (1881 - 1971)           | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Ura tai opiskelu liittyy paikkaan |
| unnianosoitukseen joka liittyy paikkaan Firenze: 'kultamitali Fiorioni e, Italia 1966'.  | Sievänen, Jaakko (1932 - 2013)        | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Kunnianosoitus liittyy paikkaan   |
| Firenze vuonna 1899.   | Silfverberg, Ida (1834 - 1899)        | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Kuollut paikassa                  |
| skelun liittyy paikassa Firenze tapahtuma: Scuola di Nudo, Firenze   | Sjöström, Wilho (1873 - 1944)         | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Ura tai opiskelu liittyy paikkaan |
| + Wladimir Swertschkoff on kuolut paikassa Firenze vuonna 1888.  | Swertschkoff, Wladimir (1821 - 1888)  | Firenze | Tapahtuma Semantisessa kansallisbiografiassa   | Kuollut paikassa                  |



**WHAT IS NEEDED  
FOR CREATING SUCH SYSTEMS?**

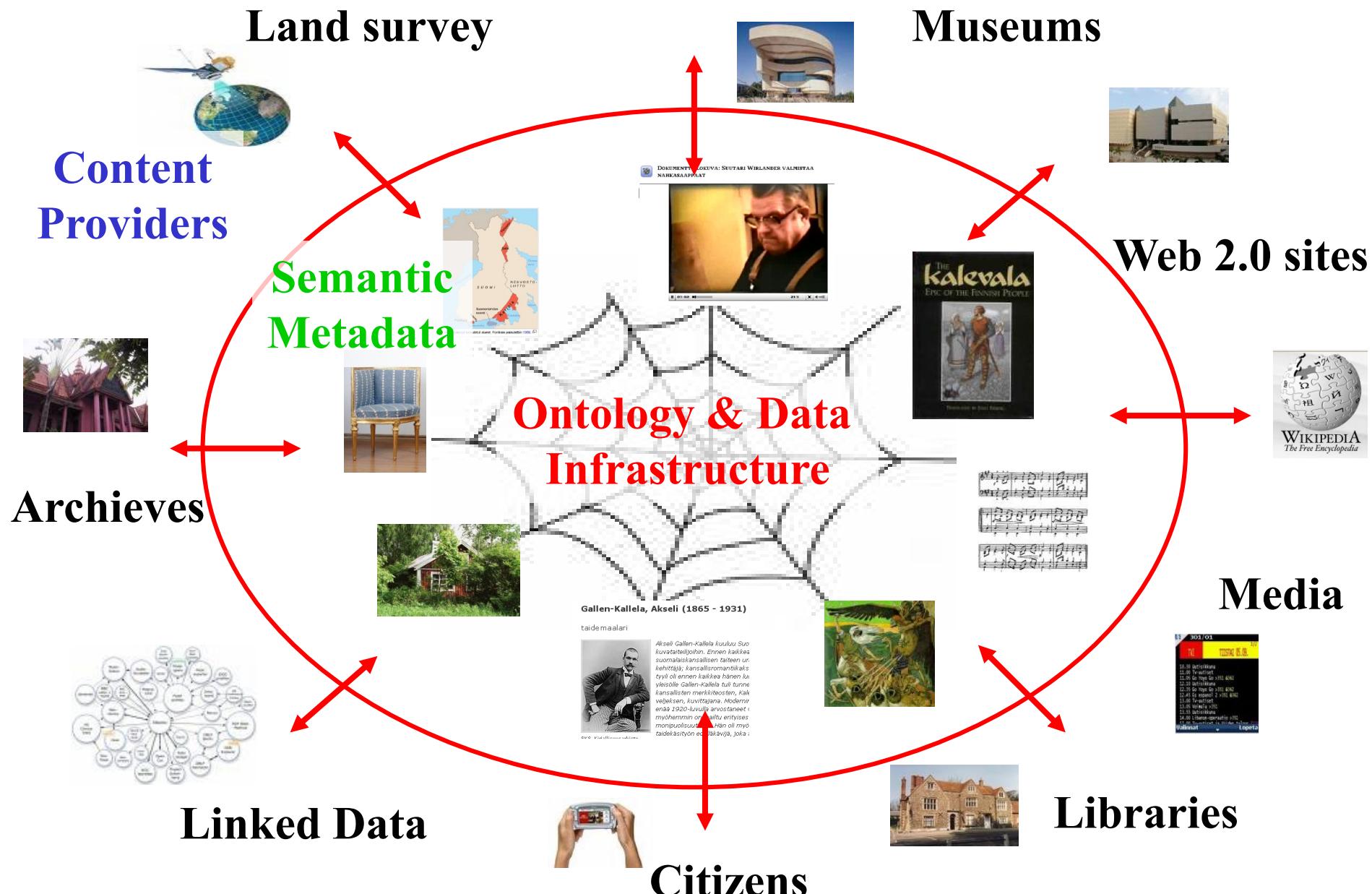
**SAMPO!**

**Sampo** = Mythical artifact of the Finnish Epic Kalevala  
that gives to its owner riches and good fortune.  
*A metaphor of technology.*



Defense of Sampo  
A. Gallen-Kallela, 1896

# "Sampo" Model for Publishing Cultural Heritage Contents: Everybody Wins!

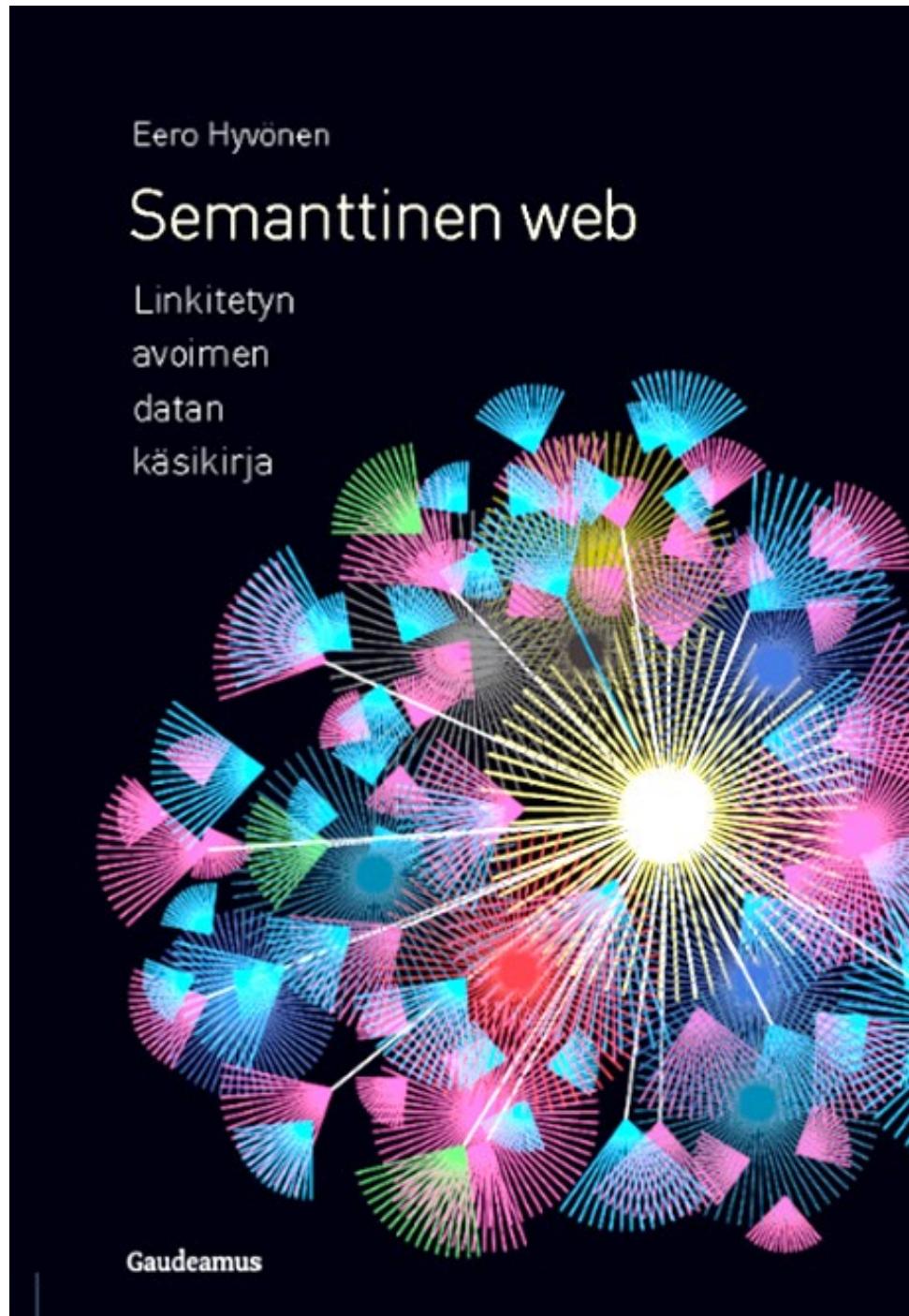


# Cultural Heritage "sampos" on the Semantic Web

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

- **CultureSampo** – Finnish Culture on the Semantic Web (2008)
- **TravelSampo** - Mobile Contextualized Services of Cultural Tourism (2011)
- **BookSampo** – Finnish Fiction Literature on the Semantic Web (2011) [2 million users in 2018]
- **WarSampo** – Finnish WW2 on the Semantic Web (2015) [230 000 users in 2018]
- **BiographySampo** - Finnish Biographies on the Semantic Web (2018) [15 000 users]
- **NameSampo** – Linked Data Workbench for Toponomastic Research (2019) [30 000 users]
  
- ***To be published in 2019-2020:***
- **WarVictimSampo 1914-1922** – National War History, with National Archives of Finland, 20.11.2019
- **MMM – Manuscript Migration Mapper**, with Oxford, Pennsylvania, IRHT (Paris), 23.11.2019
- **FindSampo** – Archaeology & Citizen Science of metal detectorists, with Finnish Heritage Agency
- **LawSampo** – Finnish Legislation and Case Law, with Ministry of Justice of Finland
- **AcademySampo** – Historical Finnish Academic People in 1640–1899

**More Info  
in Finnish:**



**More Info  
in English:**

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

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



**HELDIG**  
Helsinki Centre for Digital Humanities