SampoSampo

Connecting Everything to Everything Else

Kaikki kaikkeen kietoutuvi

Eero Hyvönen, Annastiiina Ahola, Petri Leskinen, Heikki Rantala, and Jouni Tuominen
Aalto University, Department of Computer Science, Finland
Semantic Computing Research Group (SeCo)
University of Helsinki, Helsinki Centre for Digital Humanities (HELDIG, HSSH), and
Geneva Graduate Institute, Center for Digital Humanities and Multilateralism

Towards a Web of Wisdom



Semantic Web is Based on Linked (Open) Data



“Learn how to see. Realize that everything connects to everything else.”

Leonardo da Vinci (1452-1519)

Sampo Series of LOD Services & Portals

2004





A. Gallen-Kallela: The Defense of the Sampo, 1896, Turku Art Museum

2026

1. **MuseumFinland** – Finnish Museums on the Semantic Web (2004 **[39 000 users]**
2. **CultureSampo** – Finnish Culture on the Semantic Web (2008) **[107 000 users]**
3. **TravelSampo** - Mobile Contextualized Services of Cultural Tourism (2011)
4. **BookSampo** – Finnish Fiction Literature on the Semantic Web (2011) **[1.6 million annual users]**
5. **WW1LOD** – World War I Linked Open Data (2014)
6. **WarSampo –** Finnish World War 2 on the Semantic Web (2015-19) **[1 200 000 users]**
7. **Norssi Alumni on the Semantic Web** – Historical person registry using LOD (2017)
8. **U.S. Congress Prosopographer –** U.S. Congress Legislators 1789-2018
9. **BiographySampo** - Finnish Biographies on the Semantic Web (2018-20) **[381 000 users]**
10. **NameSampo** – Linked Data Workbench for Toponomastic Research (2019) **[55 000 users]**
11. **WarVictimSampo 1914-1922** – National War History **[80 000 users]**
12. **Mapping Manuscript Migrations (MMM)** – medieval manuscripts (2020) **[9100 users]**
13. **AcademySampo –** Finnish Academic People 1640 – 1899 (2021) **[26 000 users]**
14. **FindSampo** – Archaeological Finds on the Semantic Web (2021) **[7 000 users]**
15. **WarMemoirSampo** – Memoirs of Finnish WW2 veterans (2021) **[4 400 users]**
16. **LetterSampo** – Early Modern Letters on the Semantic Web (2022)
17. **ParliamentSampo** – Parliament of Finland on the Semantic Web (2023)
18. **LawSampo –** Finnish Legislation and Case Law on the Semantic Web (2023)
19. **BookSampo II** – Semantic Search, Browsing, and Data-analyses (2023)
20. **OperaSampo** – Opera and music theater performances in Finland 1830-1960 (2023)
21. **ConfermentSampo** – Conferment Ceremonies of the University of Helsinki 1640-1899 (2023)
22. **CoinSampo** – Finnish Numismatic Public Finds 2013 – 2023 (2024)
23. **LetterSampo Finland – Finnish 19th-century Letters on the Semantic Web** (2025)
24. **SampoSampo – Connectring Everything to Everything Else** (2026)
25. **League of Nations (1920-1946) Sampo (2026)**
26. **Under constructions: ArtSampo**, **NomismaSampo**, **PASampo**, …



Sampo-UI used

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

Eero Hyvönen: **Digital Humanities on the Semantic Web: Sampo Model and Portal Series**. *Semantic Web*, 14(4), 729-744, 2023.



Legislation

Placenames & Geography

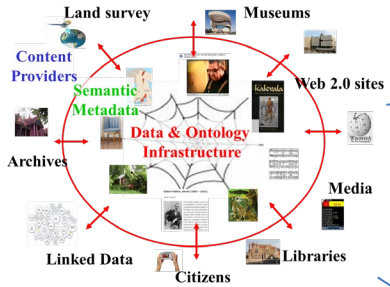
Parliamentary Speeches

Museum Artefacts

Sampo model = Swiss knife for DH

Archaeological Collections

Folklore & Narratives



Academic Traditions

Art

Biographies & Prosopography

Letters & Epistolary Data

Books & Publications

Military History

Music & Performances

Sampo Knowledge Graph Cloud

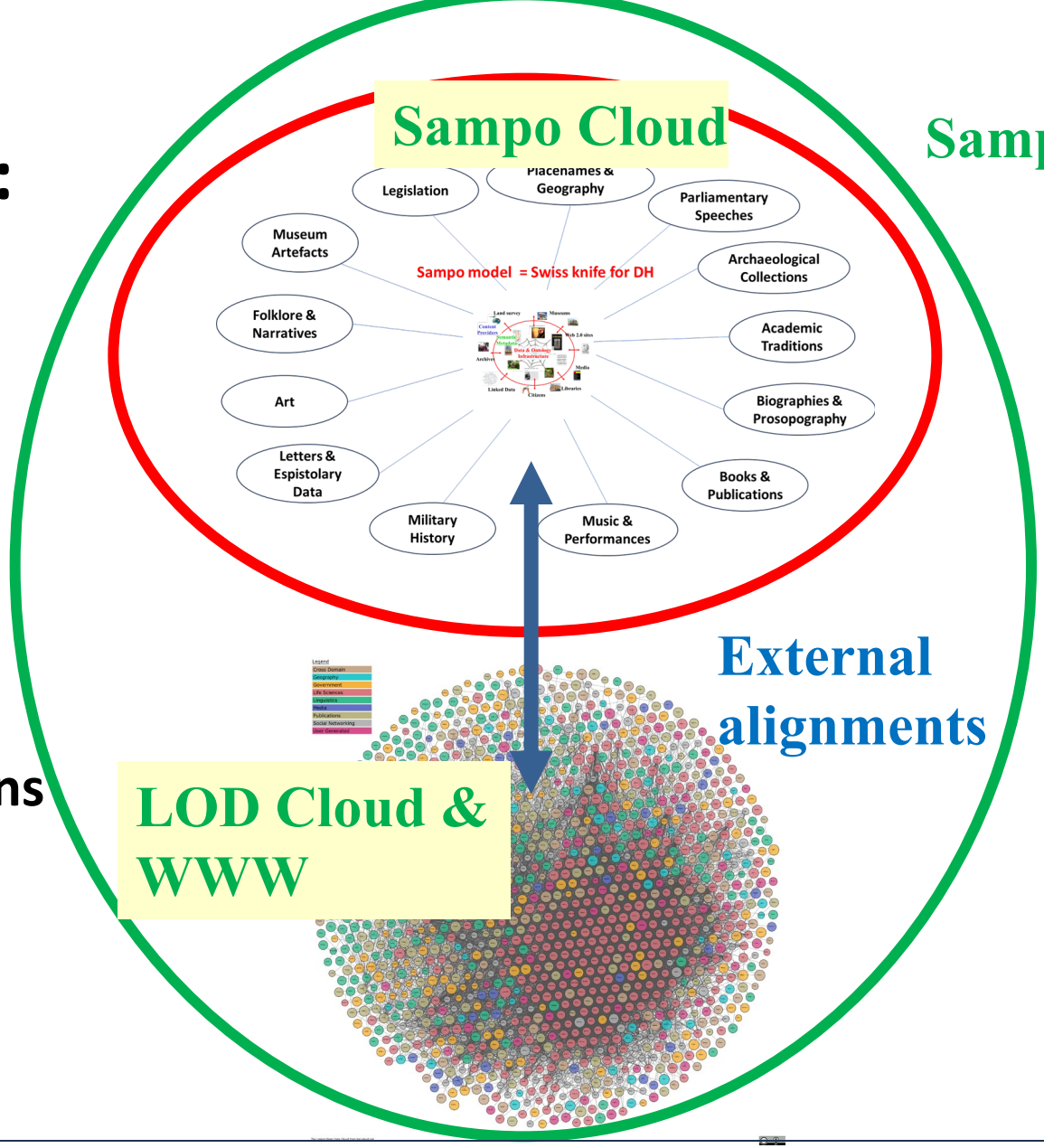


[Nano Banana 2, 2026]

SampoSampo: The Idea

Linking:

100 000 people
55 000 organizations
27 000 places





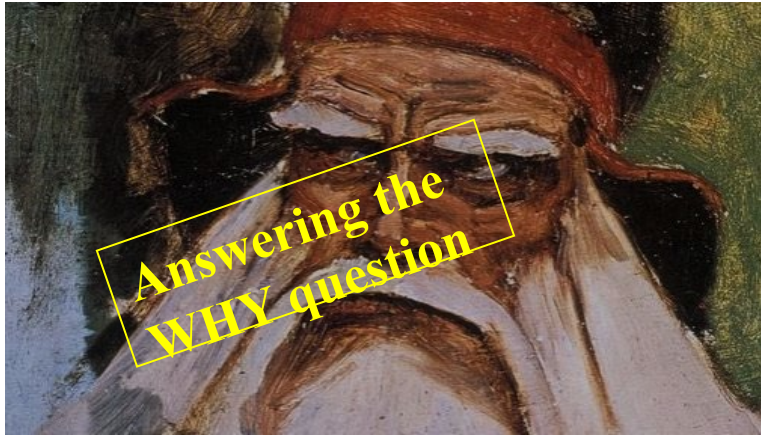


User groups:

- DH Researchers
- General Public
- IT Developers

Eero Hyvönen, Annastiina Ahola, Petri Leskinen and Jouni Tuominen: **SampoSampo: A Portal for Studying Enriched Data and Semantic Connections on a Cultural Heritage Linked Open Data Cloud**. *The Semantic Web: ESWC 2025 Satellite Events*, pp. 67-74, Springer-Verlag, 2025. pdf link

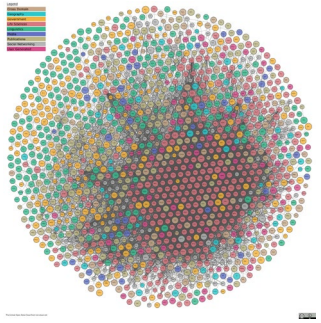
Why Is SampoSampo Needed?
New Application Ideas



1. Linking Service for creating new applications
2. For global search & browsing & data analysis across Sampos
3. Checking correctness of data between datasets
4. Knowledge discovery of connections with explanations

1. Linking Service for Creating New Applications

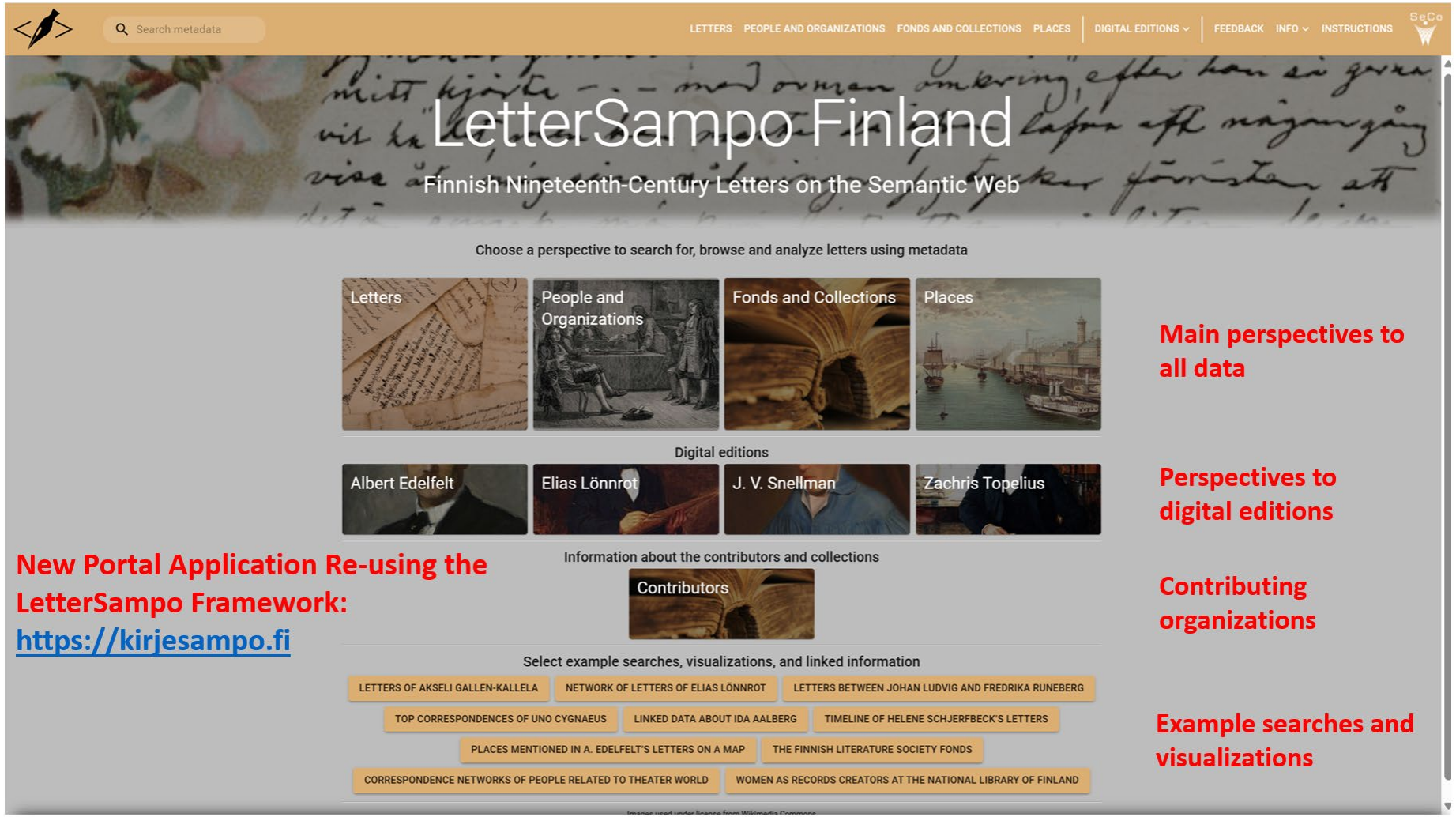
- Our idea resembles **VIAF: The Virtual International Authority File**
  - Database that aligns named entities
    - Tens national library collections
    - External datasets: Wikidata, ISNI etc.
  - Based on Linked Data
  - https://viaf.org
    - Search engine
    - Data dumps for download
- LOD Cloud:
  - https://lod-cloud.net/
- Linked Open Vobularies
  - https://datos.gob.es/en/



VIAF Contributors

- BIBSYS
- DBC (Danish Bibliographic Center)
- Flemish Public Libraries
- ISNI
- Lebanese National Library
- Library of Alexandria, Egypt
- NII (Japan)
- NUKAT Center of Warsaw University Library
- National Diet Library, Japan
- National Library and Archives of Québec
- National Library of Australia
- National Library of Catalonia
- National Library of Estonia
- National Library of Greece
- National Library of Israel
- National Library of Latvia
- National Library of Luxembourg
- National Library of Norway
- National Library of Portugal
- National Library of Spain
- National Library of the Czech Republic
- National Széchényi Library, Hungary
- National and University Library of Iceland (NULI)
- RERO - Library Network of Western Switzerland
- Slovak National Library
- Swiss National Library
- Union List of Artist Names [Getty Research Institute]
- Vatican Library
- Central Institute for the Union Catalogue of the Italian libraries
- FAST Subjects
- German National Library
- International Inventory of Musical Sources (RISM)
- Library and Archives Canada
- Library of Congress/NACO
- NUK/COBISS.SI, Slovenia
- National Central Library, Taiwan
- National Library Board, Singapore
- National Library of Argentina
- National Library of Brazil
- National Library of Chile
- National Library of France
- National Library of Ireland
- National Library of Korea
- National Library of Lithuania
- National Library of Morocco
- National Library of Poland
- National Library of Russia
- National Library of Sweden
- National Library of the Netherlands
- National and University Library in Zagreb
- Perseus
- Repertoire International de Litterature Musicale
- Sudoc [ABES], France
- Syriac Reference Portal
- United Arab Emirates University
- Wikidata

Example of Using SampoSampo for Data Enrichment: LetterSampo Finland: https://kirjesampo.fi



New Portal Application Re-using the LetterSampo Framework: https://kirjesampo.fi

Main perspectives to all data

Perspectives to digital editions

Contributing organizations

Example searches and visualizations



2. For global search & browsing & data analysis across Sampos

- Faceted search of entities as usual in Sampos but across linked datasets
- Enriched descriptions of entities with more linking
- Data analysis of enriched content in a KG cloud

Landing page:
https://samposampo.fi

Perspectives

1. People
2. Organizations
3. Places
4. Relations
5. Data sources

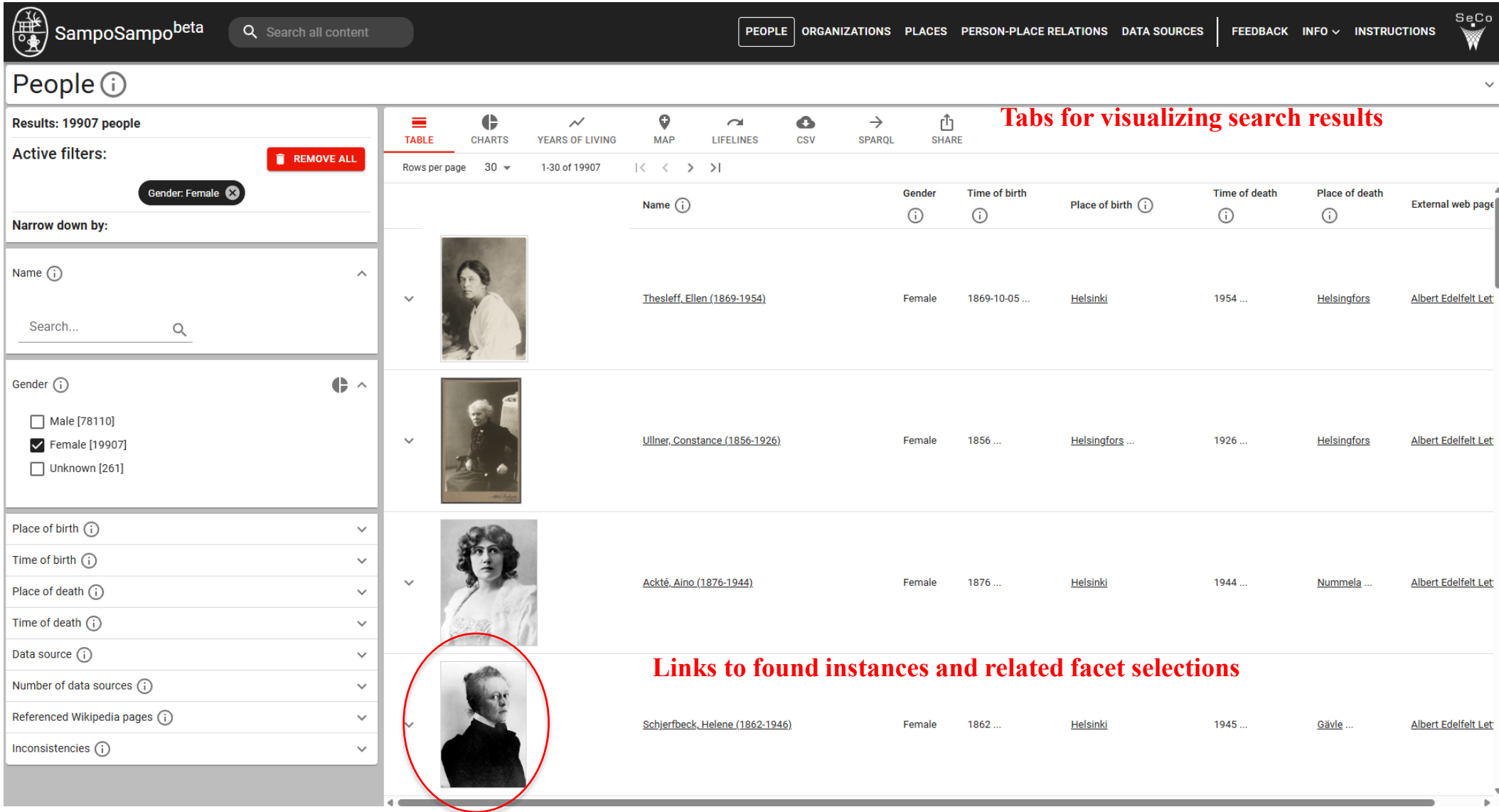
Search & analysis examples



Person Search: ca. 20 000 women found

10 search facets:

1 Name
2 Gender
3 Place of Birth
4 Time of Birth
5 Place of Death
6 Time of Death
7 Data source
8 Number of data sources
9 Reference Wikipedia pages
10 Inconsistences



Person Instance Page

Metadata elements (properties)

SampoSampo^beta

Search all content

Person

Schjerfbeck, Helene (1862-1946)

TABLE | WIKIPEDIA EXTRACTS | MAP | EXPORT

Tabs for visusolizations and data export



Image

URI http://ldf.fi/sampo/people/p6338708913222183972

Name Schjerfbeck, Helene (1862-1946)

Enriched property values from difference data sources

Preferred labels per proxy
- Scherfbeck, Helene LetterSampo Finland
- Schjerfbeck, Helena Sofia Geneanet, Wikipedia
- Schjerfbeck, Helene ArtSampo, BiographySampo, BookSampo, Edelfelt, Geneanet, KANTO, Åbo Akademi
- Schjerfbeck, Helene Sofia Geneanet
- Schjerfbeck, Helene. ULAN, Wikidata

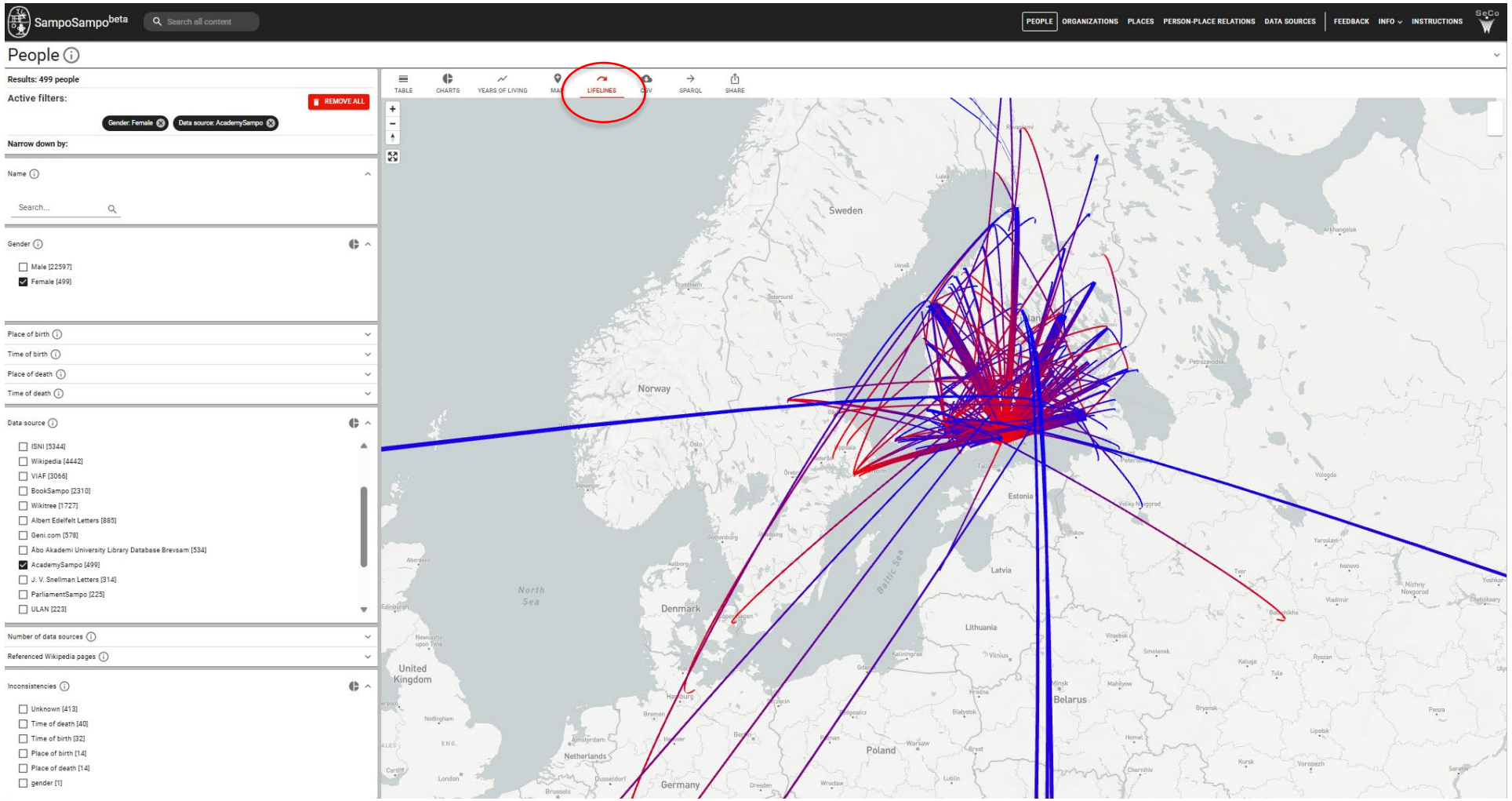
Alternative labels per proxy
- Scherfbeck, Helene Wikidata
- Schierfbeck, Helene ULAN
- Schierfberk, Helena Sofia ULAN
- Schjerfbeck, H. Wikidata
- Schjerfbeck, H. S. Wikidata
- Schjerfbeck, Helen Wikidata
- Schjerfbeck, Helena KANTO, LetterSampo Finland, Wikidata
- Schjerfbeck, Helena Sofia BiographySampo, ULAN, Wikidata

Gender Female ArtSampo, BiographySampo, BookSampo, Edelfelt, Geneanet, Geneanet, Geneanet, KANTO, LetterSampo Finland, ULAN, Wikidata, Wikipedia, Åbo Akademi

Time of birth
- 1862 ArtSampo, Edelfelt, Geneanet, KANTO, LetterSampo Finland, ULAN, Åbo Akademi
- 1862-07-10 BiographySampo, BookSampo, Geneanet, Geneanet, Wikidata, Wikipedia

Place of birth Helsingfors ArtSampo, BiographySampo, Geneanet, Geneanet, Wikidata

Person Perspective: lifelines of women in data source AcademySampo. Inconsistencies facet tells e.g. that there are 40 problems in Time of Death



3. Checking correctness of data between datasets: People

- Errors
  - ACADEMYSAMPO: von Knorring, Karl Henrik (1793–1863)
  - KANTO: Knorring, Carl Henrik von, 1973-1863
- Contradictions
  - Wasenius, Gustaf Valfrid, (1846–1928), error in birth date: 3 years
  - 1846: Dataset.ACADEMYSAMPO
  - 1848: Dataset.BIOGRAPHYSAMPO
  - 1846/1848/1849: Dataset.WIKIDATA
  - 1849: Dataset.EDELFELT
  - 1848: Dataset.KANTO
- (Possible) Duplicates
  - Westzynthius, Ericus Magni, 1744-1787 (KANTO)
  - Westzynthius, Erik, 1743-1787 (KANTO) (different father)

  - Schaumann, Anna, 1859-1928 (KANTO)
  - Schauman, A., 1859-1928 (KANTO) (same occupation)

Person Instance Page with Issues

Inconsistencies in red with provenience

Person

Agricola, Mikael (noin 1510-1557)

TABLE | WIKIPEDIA EXTRACTS | MAP | EXPORT

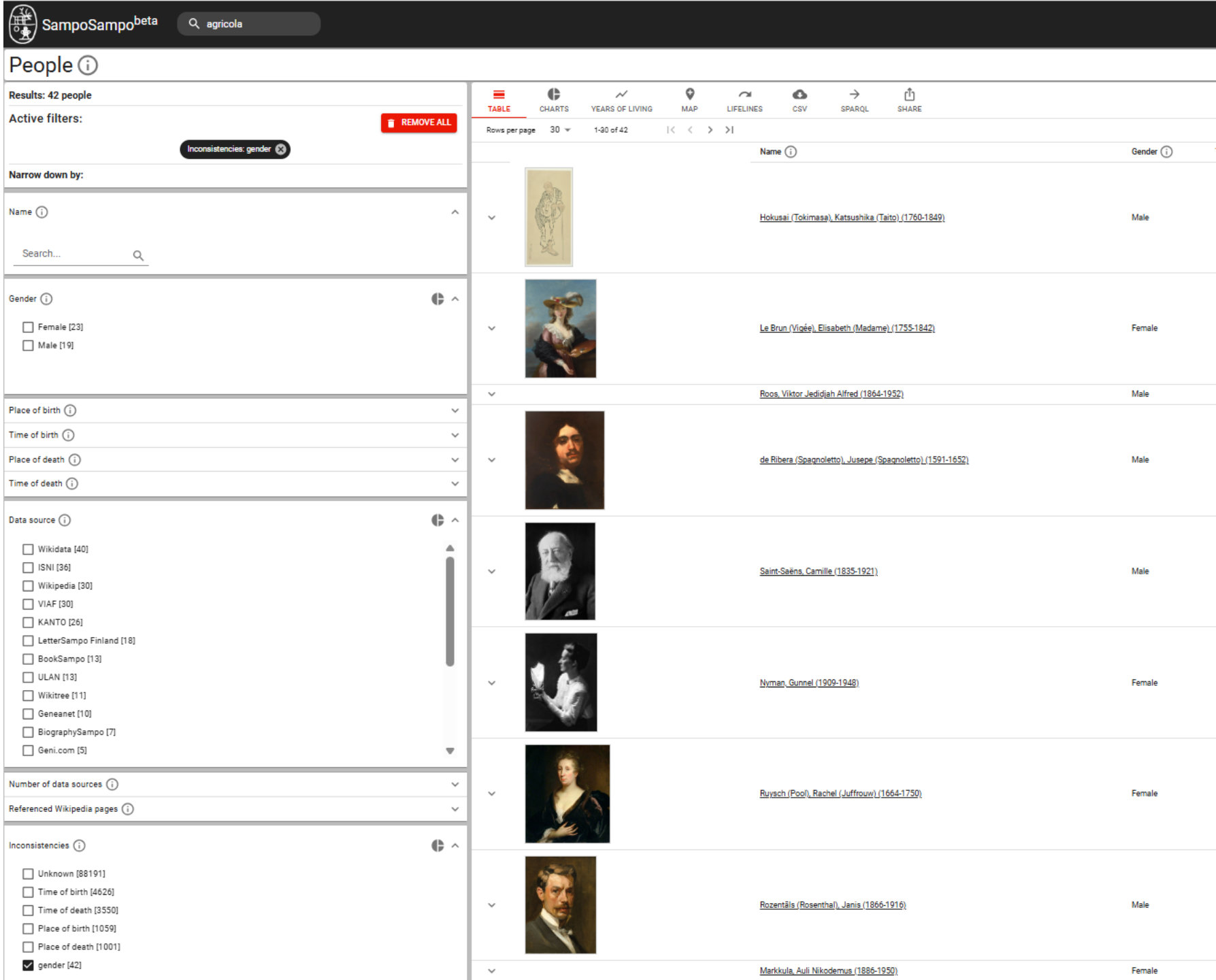


Image	
URI	http://ldf.fi/sampo/people/p5571210016405644958
Name	Agricola, Mikael (noin 1510-1557)
Preferred labels per proxy	• Agricola, Michael Olai BiographySampo • Agricola, Mikael BookSampo, LetterSampo Finland, Wikipedia • Olai, Michael Wikidata
Alternative labels per proxy	• Agricola, Michael Wikidata • Agricola, Michael Olavi Wikidata • Agricola, Mikael BiographySampo, Wikidata • Olavinpoika, Mikael Wikidata
Gender	Male BiographySampo, BookSampo, LetterSampo Finland, Wikidata, Wikipedia
Time of birth	• 1509-12-22 Wikidata • 1510 BiographySampo, BookSampo, LetterSampo Finland, Wikipedia
Place of birth	Pernå BiographySampo, BookSampo, Wikidata
Time of death	• 1557 BiographySampo, LetterSampo Finland • 1557-03-30 Wikidata • 1557-04-09 BookSampo, Wikipedia
Place of death	• Kalanti BookSampo • Nykyrka BiographySampo, Wikidata

Searching inconsistences is possible: here gender issues (42 found)

By data source

By inconsistency type



4. Knowledge discovery of connections with explanations

- Relational search by using link chains between entities
  - E.g., how persons are related to places
- Using Generative AI (LLMs) for reasoning relations
  - Prompting based of knowledge graph data and data model
  - Extending relational search over whole WWW
- Example of an explained relation: Painters -> Italy ?
  - “Elin Danielsson-Gambogi got the Florence city award in 1899”

Eero Hyvönen: **Serendipitous knowledge discovery on the Web of Wisdom based on searching and explaining interesting relations in knowledge graphs**. Journal of Web Semantics, vol. 85, Elsevier, 2025. DOI: 10.1016/j.websem.2024.100852. pdf

Annastiina Ahola, Petri Leskinen, Heikki Rantala, Jouni Tuominen and Eero Hyvönen: **Using Large Language Models for searching explainable relations in a cloud of Cultural Heritage knowledge graphs: SampoSampo as a neuro-symbolic system** Proceedings of DHNB 2026, https://doi.org/10.5617/dhnbpub.13207

A Case Study Using BiographySampo Data: Extracting New Explained Relations

- **Connections between a person and a place**
- **Not all connections are interesting**
  - Search focusing to certain types of relations
- **A KG of connections is generated (rule based and LLM)**
- **Relations are used in the SampoSampo portal UI**
  - Users can search and analyze realtions with faceted search

Annastiina Ahola, Petri Leskinen, Heikki Rantala, Jouni Tuominen and Eero Hyvönen: **Using Large Language Models for searching explainable relations in a cloud of Cultural Heritage knowledge graphs: SampoSampo as a neuro-symbolic system**
Proceedings of DHNB 2026, https://doi.org/10.5617/dhnbpub.13207

New Relations Using Extracted Named Entities

- Biographies have been run through a NER + NEL process: **people**, **places** and **organizations** mentioned in the biographies have been automatically recognized and linked to entities within BiographySampo KG
  → Able to serve as a basis for generating new person–place relations

Recognized places in the bio of Mr. Arto Paasilinna indicate possible person–place relations



KONTEKSTUAALINEN LUKIJA

Alla olevasta tekstistä on tunnistettu henkilöt, paikat, ja organisaatiot automaattisin menetelmin.

Arto Paasilinna on lukijoiden laajan suosion saavuttanut kirjailija, jonka teokset ovat menestyneet myös ulkomailla. Kirjallisuuskriitikoita hänen teoksensa eivät ole onnistuneet vakuuttamaan. Paasilinna on kirjailijana humoristi, ja hänen romaaneidensa aiheet ovat olleet varsin mielikuvituksellisia.

Arto Paasilinna on sodan ja pohjoisen poika. Hänen isänsä työskenteli poliisina Petsamossa, ja Arto syntyi perheen sotapakolaisuuden aikana Kittilässä. Paasilinnan veljessarja on suomalainen menestystarina: Erno Paasilinna oli toisen maailmansodan jälkeisten vuosien kiitetyimpiä, kiistellyimpiä ja palkituimpia kirjailijoita ja Reino Paasilinna aiemmin kohuttu televisiotoimittaja, myöhemmin diplomaatti ja poliitikko. Kirjailijana suurimman kansansuosion on saavuttanut Arto Paasilinna. Hänen tuotantonsa on laaja, valtaosa romaaneja. Lukijoiden suosikkina ja painosten "herrana" Paasilinna sijoittuu Kalle Päätalon ja Laila Hietamiehen (vuodesta 2005 Hirvisaari) rinnalle.

https://biografiasampo.fi/henkilo/p4752

Generating Person–place Relations Using RAG

- The model used was OpenAI's **GPT-5 mini** (web searches tool enabled)
- The mentions were used for prompting for finding relational data with explanations.
- The model was also asked to list the internet sources used and to provide a brief explanation of what information the model claims to have obtained from each source.
- Such **provenance data** can be used to assess the reliability of the relation
- The model's responses were requested in **JSON format**, which was subsequently converted to **RDF format** for interoperability with other relations included in SampoSampo

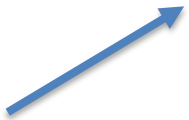
Evaluation (Based on a Small Sample)

- **Relations**: whether the nodes of the relation match the ones stated in the relation description and whether the relation description is a true statement
  - semantic *interestingness* was not evaluated
- **Statements**: whether the specified source contains all the information mentioned in the statement. If there was extra information, the statement was considered partially correct.

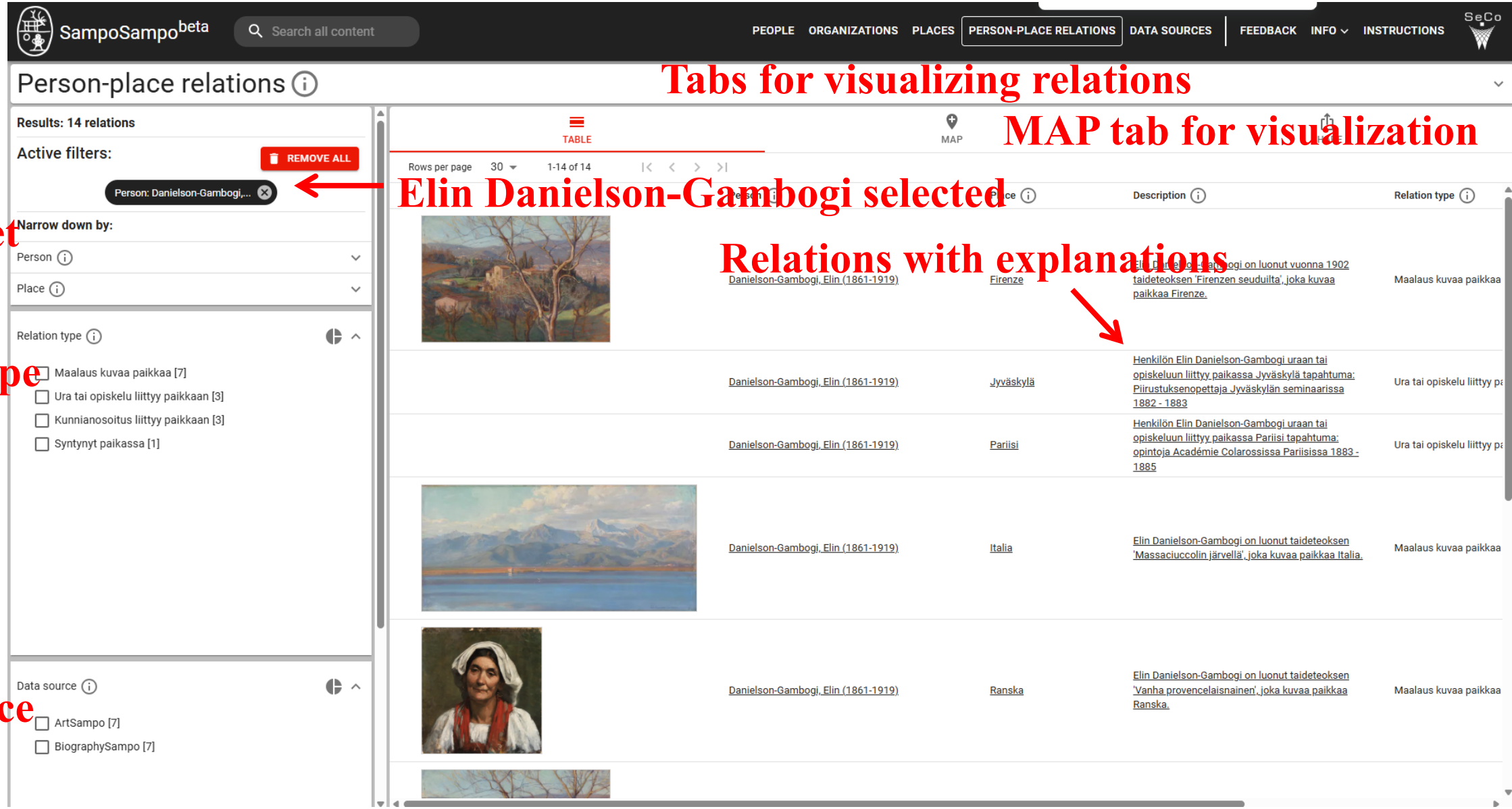
Type	Total count	Correct	Partially correct	Incorrect
Relations	50	50	0	0
Statements	123	111	10	2

- **Serendipity:** Interestingly, LLMs found new relation types that were not the original KG and that we did not anticipate!

Relation type	Count
Existing relation types	*37*
– Place of death	6
– Place of birth	4
– Career or studies related to place	20
– Event took place at	4
– Received honor related to place	3
New relation types	*13*
– Received honor in place	2
– Educated at place	1
– Resided in place	2
– Membership in institution in place	1
– Family member worked in place	1
– Leisure activity in place	1
– Archives located in place	1
– Stage design depicts place	1
– Citizen of place	1
– Held local government position in place	1
– Held local church office in place	1



Faceted Search for Person-place Relations



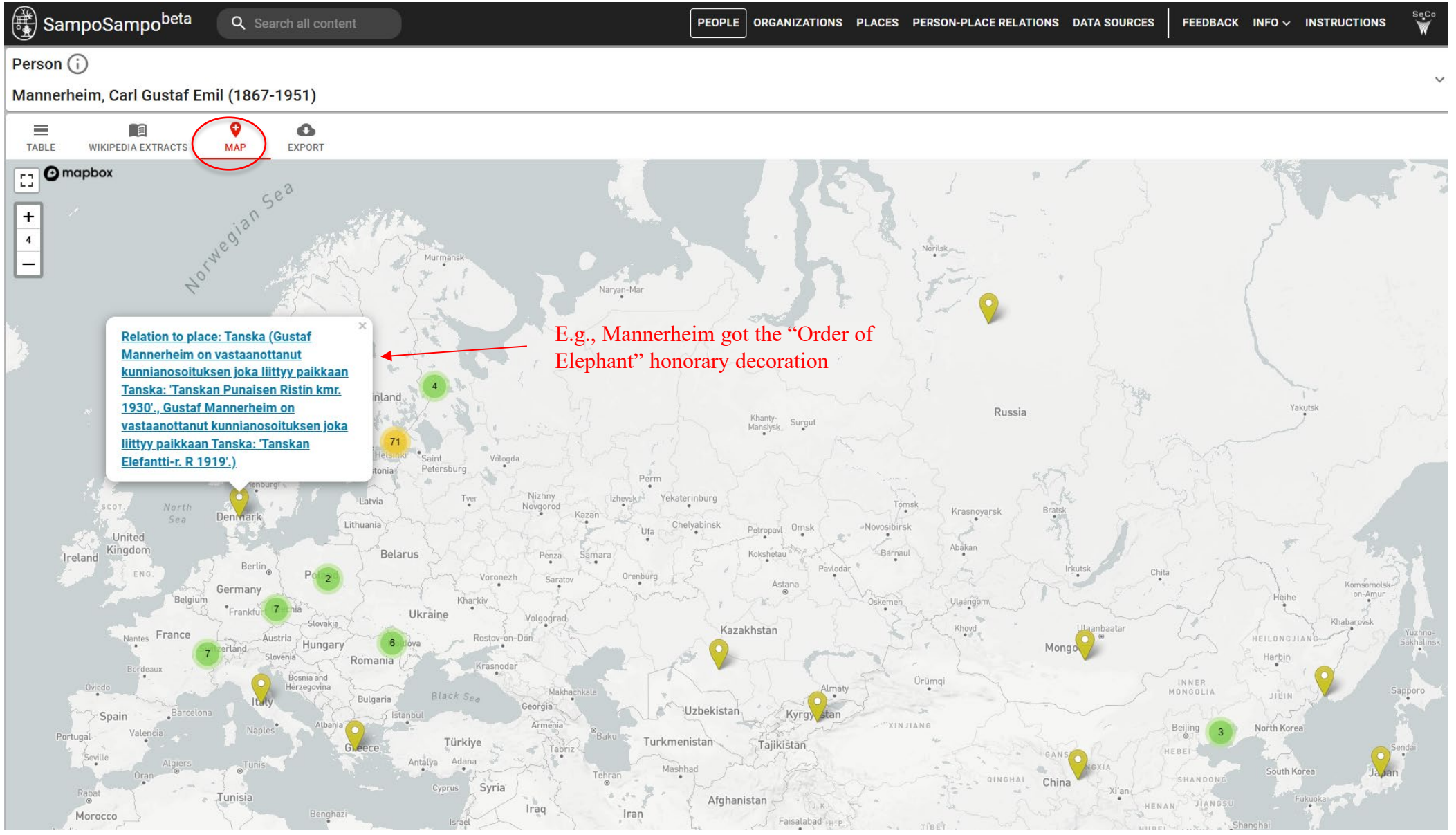
Person facet

Place facet

Relation type facet

Data source facet

How is C. G. E Mannerheim Related to Places?



LESSONS LEARNED 2002-2026

1. Develop shared LOD infra **step-by-step together**
2. Reuse LOD infra: **easy to develop/use applications** on top of it
3. LOD is very **resilient for different application domains**
4. **LLM:s open a promising way to neuro-symbolic AI**



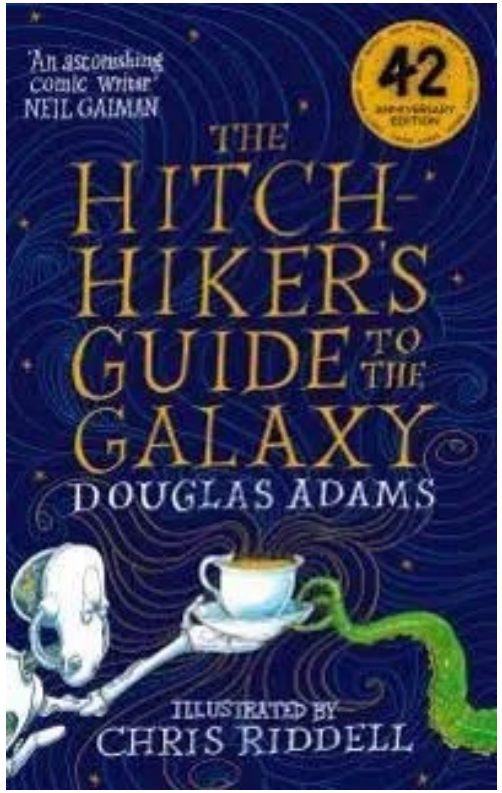
Our work is today part of the National Finnish FIN-CLARIAH/DARIAH-FI Research Infrastructure

https://seco.cs.aalto.fi/projects/fin-clariah/

Funded by the European Union
NextGenerationEU

Suomen Akatemia
Finlands Akademi
Research Council of Fin

What is the Meaning of Life,
the Universe, and Everything?
“42”



Explainable AI is needed for humanities research,
and symbolic LOD and neuro-symbolic AI is a way to go!

More Information – Questions?

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

