

Kirjesampo osana kansallista FIN-CLARIAH/DARIAH- FI-tietoinfrastruktuuria: linkitetty avoin data, datapalvelut ja rajapintojen käyttö

**Kirjesampo-julkistustilaisuus
Suomalaisen Kirjallisuuden Seura
27.5.2025**

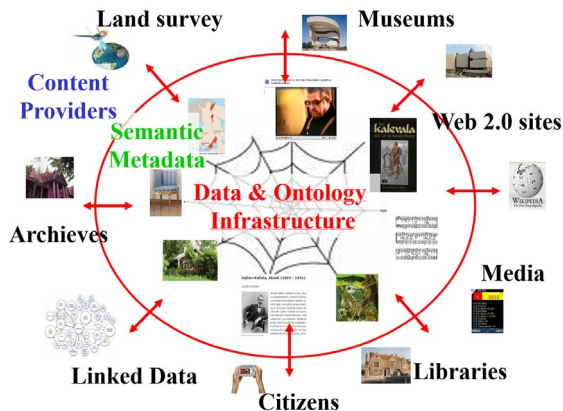
Jouni Tuominen, Helsingin yliopisto

Sampo Model for 1) Human Users and 2) Machines



LINKED DATA SERVICE

Findable, Accessible, Interoperable, Re-usable
FAIR data!



Linked Data Finland
Living Laboratory Data Service for the Semantic Web

<https://ldf.fi/>

APPLICATIONS & TOOLS

Use Cases

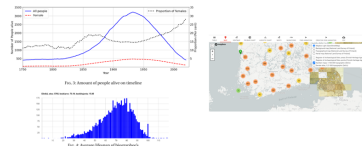
Sampo Portals & Applications



Samos:

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

Digital Humanities Research



Tools

Sampo-UI, etc.

Google Colab, Jupyter,
YASGUI, etc.



{REST API}

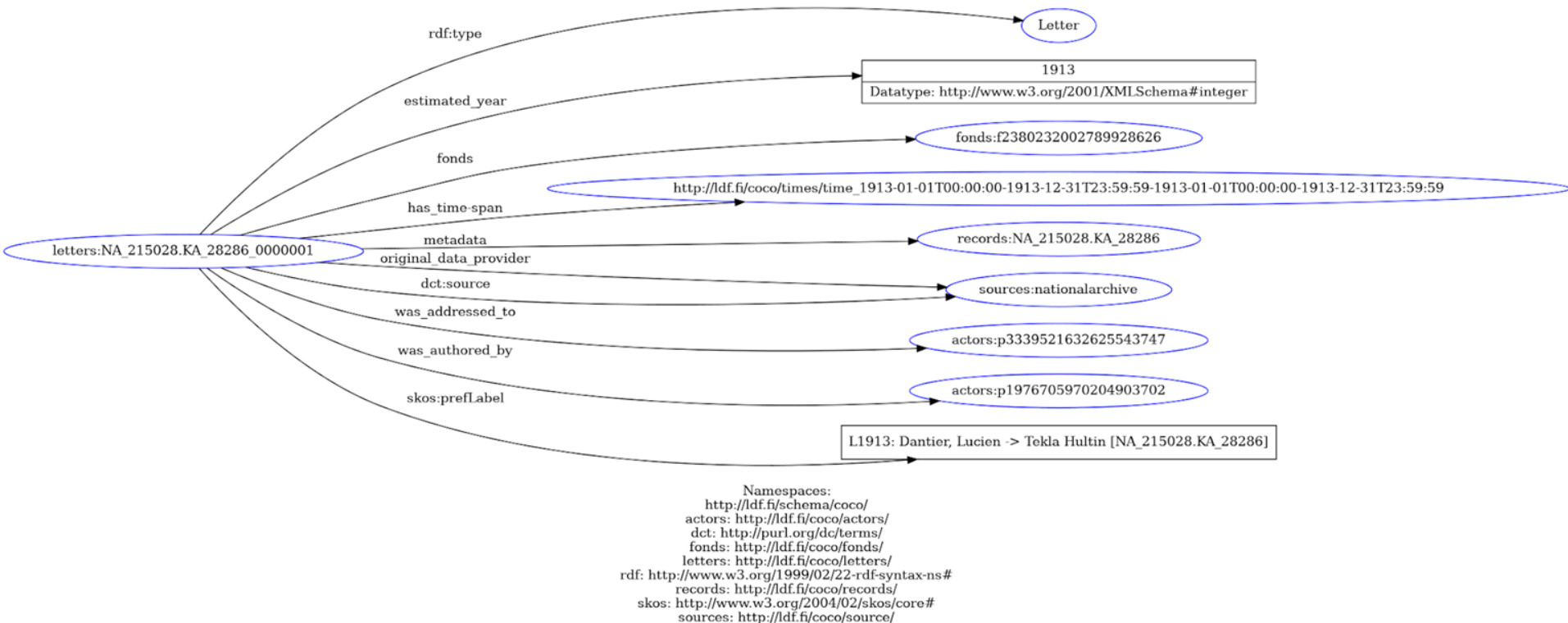
References

Eero Hyvönen: **How to Create a National Cross-domain Ontology and Linked Data Infrastructure and Use It on the Semantic Web**. *Semantic Web*, 2024. DOI: 10.3233/SW-243468.

Eero Hyvönen: **Digital Humanities on the Semantic Web: Sampo Model and Portal Series**. *Semantic Web*, vol. 14, no. 4, pp. 729-744, 2023.

Esko Ikkala, Eero Hyvönen, Heikki Rantala and Mikko Koho: **Sampo-UI: A Full Stack JavaScript Framework for Developing Semantic Portal User Interfaces**. *Semantic Web*, vol. 13, no. 1, 69-84, 2022.

Graph-based data model (RDF) (try!)



Linked Data Finland Living Lab <http://ldf.fi>



[Home](#)
[Project](#)
[Datasets](#)
[Search Data](#)
[Schemas](#)
[Services](#)
[Policies](#)
[Documentation](#)
[Validation](#)
[Linked Data Science](#)
[Applications](#)
[Your Data?](#)
[Linked Data School](#)

Linked Data Finland

Living Laboratory Data Service for the Semantic Web

This site is the Living Laboratory of the [FIN-CLARIAH](#) research infrastructure initiative, conducted by the [Semantic Computing Research Group](#) at [Aalto University](#) in collaboration with University of Helsinki and a large consortium of other Finnish universities and organizations.

Our goal is to make life easier for both Linked Data publishers and consumers on the Web. We base our work on the Linked Data principles and stack of standards founded on the semantic data model [RDF](#), standardized query language [SPARQL](#), and [live HTTP URIs](#) for identifiers.

5-star Linked Data

The baseline of our work is the [5-star Linked Data model](#), proposed by Tim Berners-Lee.

- ★ Make data available on the Web in whatever format.
- ★★ Make data available as structured data (e.g., Excel instead of an image scan of a table).
- ★★★ Use non-proprietary formats (e.g., CSV instead of Excel format).
- ★★★★ Use URIs to denote things, so that people can point at your data.
- ★★★★★ Link your data to other data to provide context.

8-star Linked Data

However, in our opinion, providing 5-star Linked Data is just the beginning. To actually make use of the datasets, consumers need more support in getting to know the underlying data models used, how well the data actually conforms to the models, as well as trustworthiness of the data. To these ends, we extend the model with three additional stars to encourage data publishers to make their data more re-usable:

- ★★★★★ Describe explicitly and publish the schemas used in the dataset (unless the schemas are already available elsewhere).
- ★★★★★ Explain the quality of the dataset against the schemas used in it, so that the user can tell whether the data quality matches her needs.
- ★★★★★ Give explanations when the data is factually correct with respect to the real world and when possibly not.

Eero Hyvönen and Jouni Tuominen: 8-star Linked Open Data Model: Extending the 5-star Model for Better Reuse, Quality, and Trust of Data. *Posters, Demos, Workshops, and Tutorials of the 20th International Conference on Semantic Systems (SEMANTICS 2024)*, vol. 3759, CEUR Workshop Proceedings, September, 2024. <https://ceur-ws.org/Vol-3759/paper4.pdf>

Metadata of >1,200,000 letters and 100,000 letter senders and recipients!



[Home](#)
[Project](#)
[Datasets](#)
[Search Data](#)
[Schemas](#)
[Services](#)
[Policies](#)
[Documentation](#)
[Validation](#)
[Linked Data Science](#)
[Applications](#)
[Your Data?](#)
[Linked Data School](#)

LetterSampo Finland – Finnish Nineteenth-Century Letters on the Semantic Web

Linked Data Finland



This dataset includes data regarding Finnish Nineteenth-Century epistolary data. The RDF data has been converted from the source data provided by multiple Finnish cultural heritage organizations; we call the new dataset LetterSampo Finland.

The Knowledge Graph contains harmonized data of:

1. Letters,
2. Actors,
3. Fonds,
4. Places,
5. Occupations.

The data model is designed for epistolary letter data, including the information about the sender, recipient, time of sending, and places of sending and receiving.

To test and demonstrate its usefulness, this Knowledge Graph is in use in the semantic portal [LetterSampo Finland](#), explained in more detail in the [project page](#).

The Knowledge Graph is published on [Zenodo](#) with a version history, and it can be downloaded as RDF files or CSV datasheets.

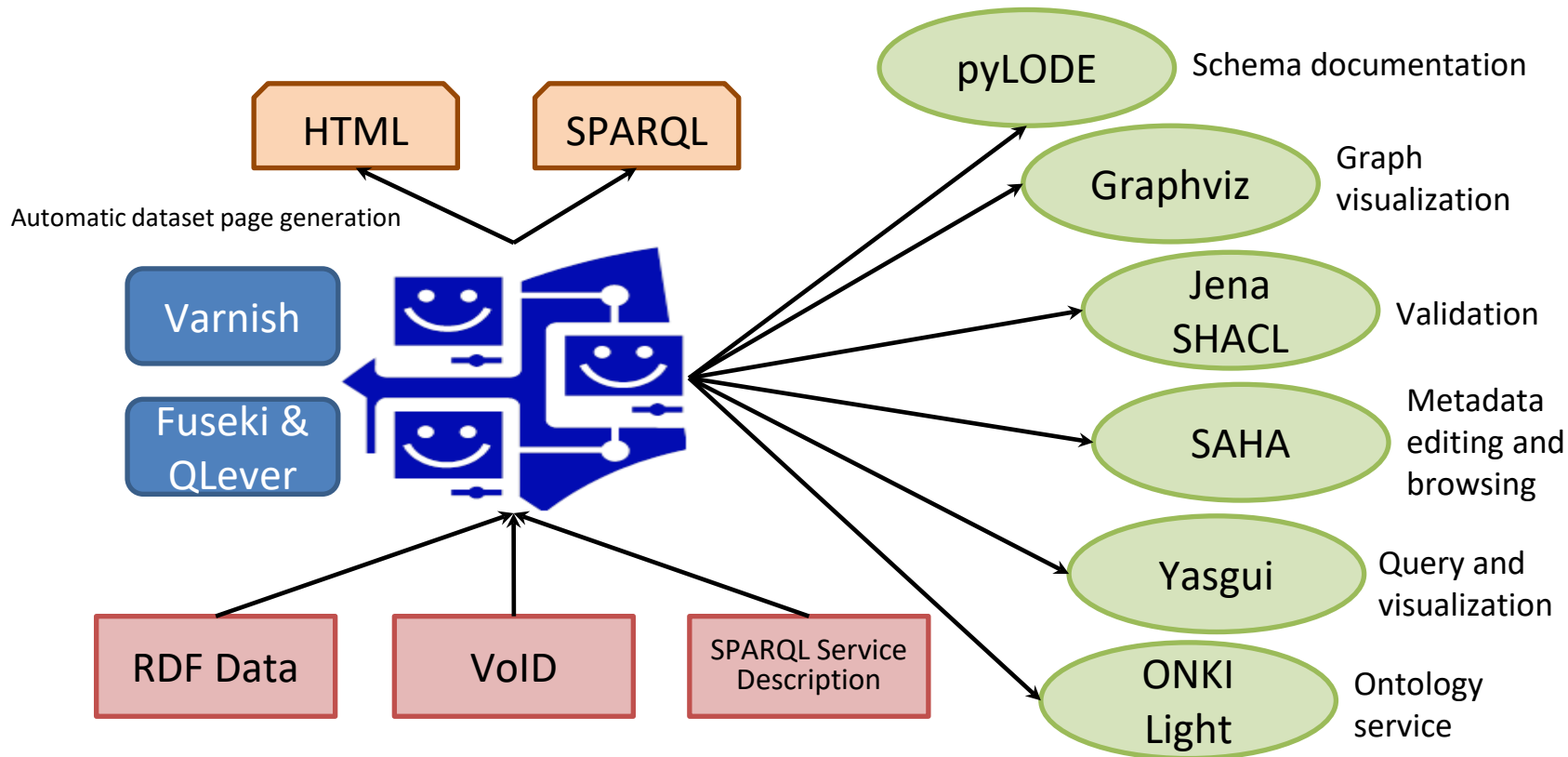
License

[Creative Commons Attribution 4.0](#)



See possible graph-specific licenses below.

LDF.fi – generic Linked Data publication platform



provided:p00000000000000000370

[illegible]

LDF.fi: SPARQL endpoints for querying data for analysis, visualization, mash-up apps (CSV, JSON, XML) (try!)

[http://fdoos/fdoos.html](#)

```

1 PREFIX dets: <http://purl.org/dc/terms/>
2 PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>
3 PREFIX : <http://ldf.fi/schema/coco/>
4
5 SELECT ?source1 ?source2 (COUNT(DISTINCT ?provided_person) as ?shared_person_amount)
6 {
7   ?provided_person a :ProvidedActor .
8   [ ?proxy_for ?provided_person ;
9     a crm:E21_Person ;
10    dct:isource ?source1 .
11   [ ?proxy_for ?provided_person ;
12     a crm:E21_Person ;
13     dct:source ?source2 .
14   FILTER (STR(?source1) > STR(?source2))

```

Press CTRL + <spacebar> to autocomplete

Table

Response

Gallery

Chart

Geo

Geo-3D

Geo events

Markup

Network

Pivot

Timeline

119 results in 8.559 seconds

Simple view

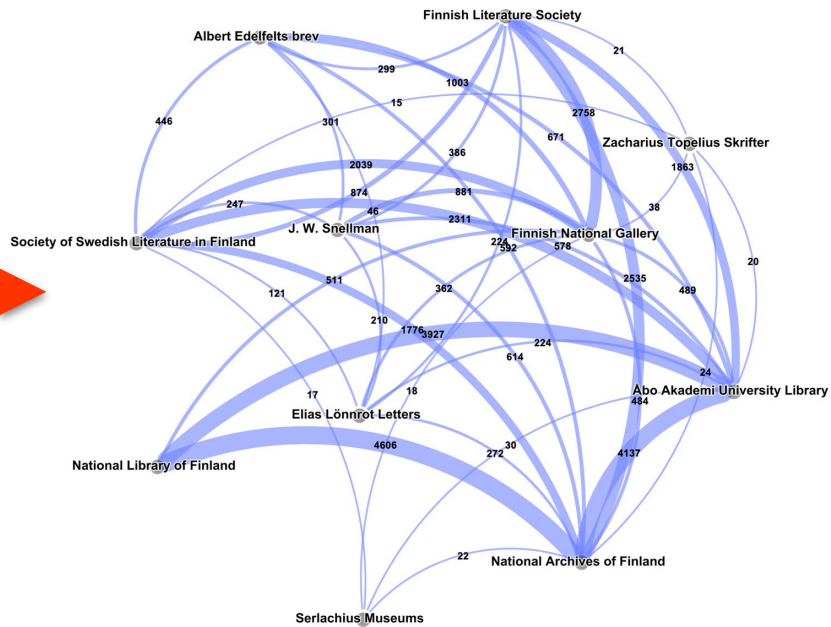
Ellipse

Filter query results

Page size:

50

source1	source2	shared_person_amount
<http://ldf.fi/coco/source/nationallibrary>	<http://ldf.fi/coco/source/nationalarchive>	"4626"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationalarchive>	<http://ldf.fi/coco/source/abokademik>	"4174"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationallibrary>	<http://ldf.fi/coco/source/abokademik>	"3990"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/sks>	<http://ldf.fi/coco/source/nationalarchive>	"2616"<http://www.w3.org/2001/XMLSchema#integer>
<https://sls.fi/>	<http://ldf.fi/coco/source/abokademik>	"2596"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/sks>	<http://ldf.fi/coco/source/nationallibrary>	"2511"<http://www.w3.org/2001/XMLSchema#integer>
<https://sls.fi/>	<http://ldf.fi/coco/source/nationalarchive>	"2027"<http://www.w3.org/2001/XMLSchema#integer>
<https://sls.fi/>	<http://ldf.fi/coco/source/nationallibrary>	"1960"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/sks>	<http://ldf.fi/coco/source/abokademik>	"1928"<http://www.w3.org/2001/XMLSchema#integer>
<https://sls.fi/>	<http://ldf.fi/coco/source/sks>	"975"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationallibrary>	<http://ldf.fi/coco/source/edelfelt>	"777"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/edelfelt>	<http://ldf.fi/coco/source/abokademik>	"671"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationalarchive>	<http://ldf.fi/coco/source/sks>	"591"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationallibrary>	<http://ldf.fi/coco/source/nationalgallery>	"511"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationalgallery>	<http://ldf.fi/coco/source/abokademik>	"498"<http://www.w3.org/2001/XMLSchema#integer>
<https://sls.fi/>	<http://ldf.fi/coco/source/edelfelt>	"494"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/nationalgallery>	<http://ldf.fi/coco/source/nationalarchive>	"491"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/sks>	<http://ldf.fi/coco/source/nationalgallery>	"331"<http://www.w3.org/2001/XMLSchema#integer>
<http://ldf.fi/coco/source/sks>	<http://ldf.fi/coco/source/nationalarchive>	"323"<http://www.w3.org/2001/XMLSchema#integer>
<https://sls.fi/>	<http://ldf.fi/coco/source/nationalgallery>	"322"<http://www.w3.org/2001/XMLSchema#integer>



LDF.fi: Download data graphs (Turtle, JSON-LD, RDF/XML, ...) via SPARQL HTTP Graph Store Protocol (try!)

```
labels:s2166837430437462088
  rdf:type      skosxl:Label;
  tmp:weight    3;
  sch:familyName "Flodin";
  sch:givenName  "Ferdinand";
  skos:prefLabel "Flodin, Ferdinand" .
```

```
labels:s1944721296601468395
  rdf:type      skosxl:Label;
  tmp:weight    1 , 4;
  sch:familyName "Gripenberg";
  sch:givenName  "Wilhelmina";
  skos:prefLabel "Gripenberg, Wilhelmina" .
```

```
labels:s1184357776282920260
  rdf:type      skosxl:Label;
  tmp:weight    2;
  :name_addition "von";
  sch:familyName "Schoultz";
  sch:givenName  "Carl";
  skos:prefLabel "von Schoultz, Carl" .
```

```
labels:s1478893002208474798
  rdf:type      skosxl:Label;
  tmp:weight    9;
  sch:familyName "Nyström";
  sch:givenName  "Gustaf";
  skos:prefLabel "Nyström, Gustaf" .
```

```
labels:s3151972063766390443
  rdf:type      skosxl:Label;
  tmp:weight    1 , 366 , 486;
  sch:familyName "Hackman";
  sch:givenName  "Wilhelm";
  skos:prefLabel "Hackman, Wilhelm" .
```

```
labels:s2375809655731754294
  rdf:type      skosxl:Label;
  tmp:weight    2;
  sch:familyName "Myréén";
  sch:givenName  "Oskar";
  skos:prefLabel "Myréén, Oskar" .
```

```
labels:s3215776551662322704
  rdf:type      skosxl:Label;
  tmp:weight    6;
  sch:familyName "Strömberg";
  sch:givenName  "Max";
  skos:prefLabel "Strömberg, Max" .
```



LetterSampo Finland

Finnish Nineteenth-Century Letters on the Semantic Web

Choose a perspective to search for, browse and analyze letters using metadata

Letters

People and
Organizations

Fonds and Collections

Places

Digital editions

Albert Edelfelt

Elias Lönnrot

J. V. Snellman

Zachris Topelius

Information about the contributors and collections

Contributors

Select example searches, visualizations, and linked information

LETTERS OF AKSELI GALLEN-KALLELA

NETWORK OF LETTERS OF ELIAS LÖNNROT

LETTERS BETWEEN JOHAN LUDVIG AND FREDRIKA RUNEBERG

TOP CORRESPONDENCES OF UNO CYGNAEUS

LINKED DATA ABOUT IDA AALBERG

TIMELINE OF HELENE SCHJERFBECK'S LETTERS

PLACES MENTIONED IN A. EDELFELT'S LETTERS ON A MAP

THE FINNISH LITERATURE SOCIETY FONDS

CORRESPONDENCE NETWORKS OF PEOPLE RELATED TO THEATER WORLD

WOMEN AS RECORDS CREATORS AT THE NATIONAL LIBRARY OF FINLAND

Images used under license from Wikimedia Commons

LetterSampo Finland data dumps (static files)

- **Zenodo**

- Linked Data format: RDF files
- Other formats: CSV files
- Versioning, digital object identifier (DOI)

Published May 27, 2025 | Version 1.0.0

Dataset  Open

LetterSampo Finland

Ahola, Annastina¹ ; Björklund, Maria² ; Drobac, Senka^{3,1} ; Enqvist, Johanna² ;
Hyvönen, Eero Antero^{3,1} ; Jauhainen, Iida² ; Kanervo, Sanni² ; Koho, Mikko^{1,3} ;
Koho, Ossi³ ; Kukkonen, Ville² ; Leal, Rafael¹ ; Leskinen, Petri^{1,3} ; La Mela, Matti⁴ ;
Paloposki, Hanna-Leena² ; Pikkanen, Ilona² ; Poikkimäki, Henna¹ ; Rantala, Heikki¹ ;
Sainio, Suvi² ; Tuominen, Jouni^{1,3} ; Wahjoe, Muhammad Faiz¹ 

Show affiliations

This dataset includes data regarding Finnish Nineteenth-Century epistolary data. The RDF data has been converted from the source data provided by multiple Finnish cultural heritage organizations; we call the new dataset LetterSampo Finland.

The Knowledge Graph contains harmonized data of:

1. Letters,
2. Actors,
3. Fonds,
4. Places,
5. Occupations.

The data model is designed for epistolary letter data, including the information about the sender, recipient, time of sending, and places of sending and receiving.

To test and demonstrate its usefulness, this Knowledge Graph is in use in the semantic portal [LetterSampo Finland](#), explained in more detail in the [project page](#).

0
VIEWS

0
DOWNLOADS

Show more details

Versions

Version 1.0.0	May 27, 2025
10.5281/zenodo.15210590	

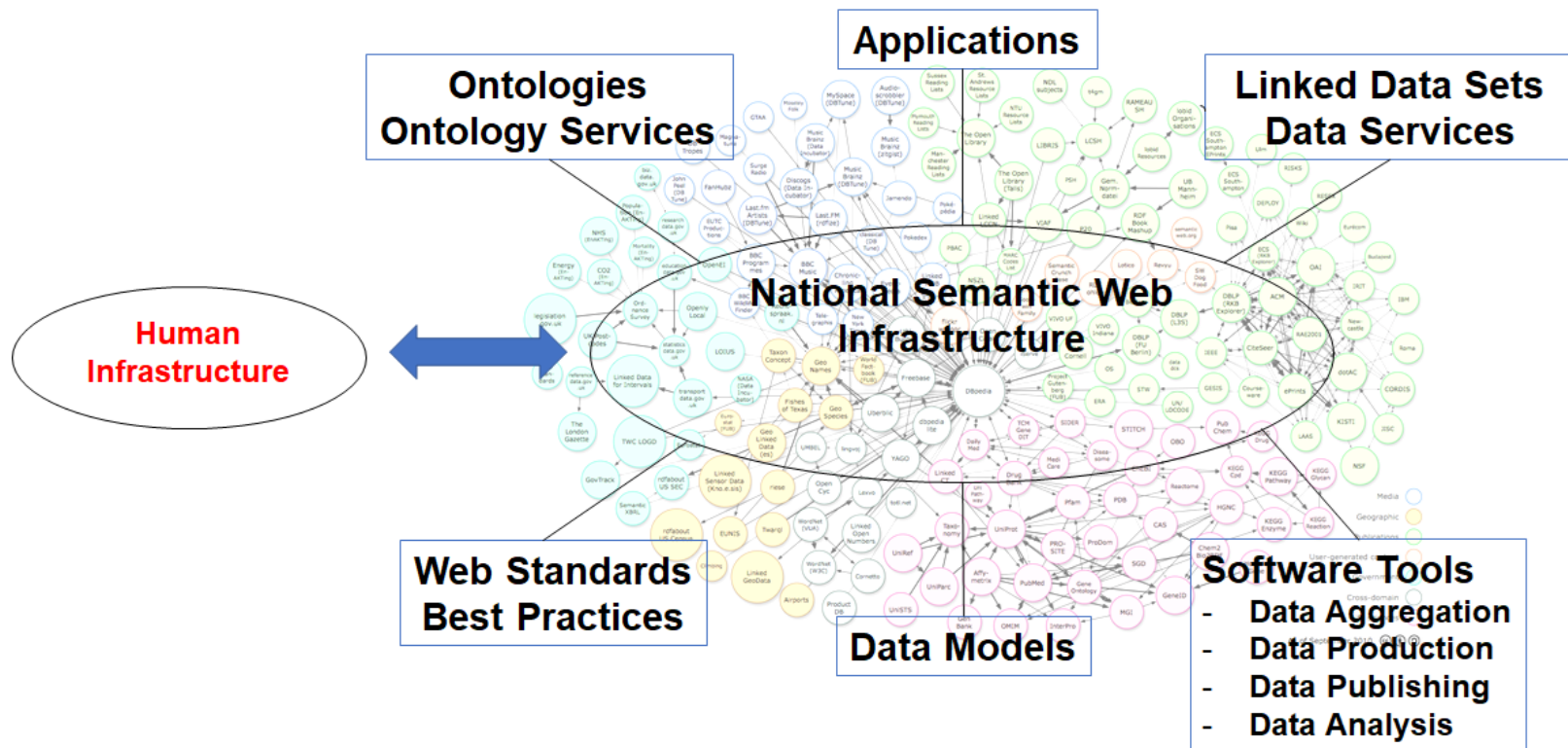
Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.15210589](#). This DOI represents all versions, and will always resolve to the latest one. [Read more](#).

External resources

Indexed in



Linked Data infrastructure



Vocabulary of Finnish Actors

- SampoSampo

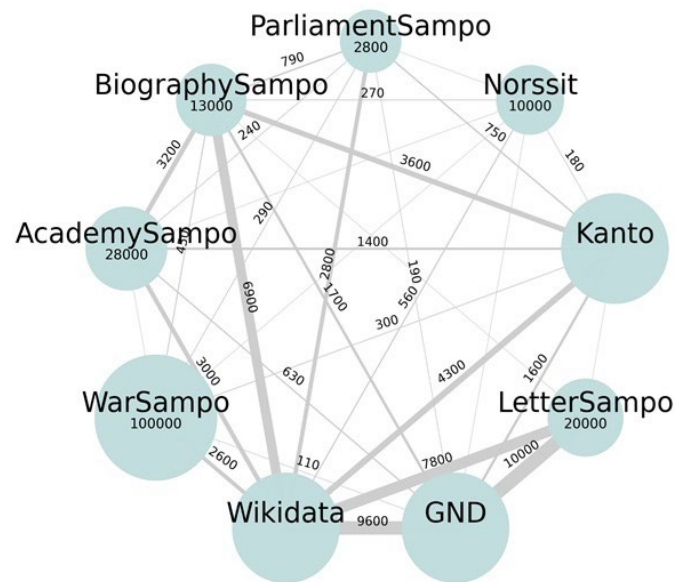
Historical Finnish People, Organizations, and Places

Biographical information about 230000 historical people related to Finland

Name variations, gender, years and places of birth and death ...

- Collected from multiple datasources:

1. AcademySampo (23000 entries)
2. BiographySampo (22500)
3. Wikidata (84000)
4. Kanto (29000)
5. BookSampo (9800)
6. ParliamentSampo (2100)
7. Norssi Highschool Alumni (2600)
8. Union List of Artist Names (1500)
9. Wikitree (12000)
10. Geneanet (144000)
Parts of CoCo data providing biographical details
11. Edelfelt Letters (3300)
12. Snellman Letters (3300)
13. Åbo Akademi University Library letters (1570)



Sampo infrastructure

- **All software components are run as Docker containers**
 - Portable services, can be run on a local machine/server
- **Triple stores (Fuseki, QLever) run on virtual machines**
- **Web applications (Varnish for routing + cache, www.ldf.fi, portals) run on OpenShift container cloud**
- **Part of the FIN-CLARIAH/DARIAH-FI research infrastructure program**
- **Technical infrastructure provided by CSC – IT Center for Science (by Ministry of Education and Culture)**

Kiitos!



Aalto University
School of Science



UNIVERSITY OF HELSINKI



SUOMALAISEN
KIRJALLISUUDEN
SEURA



HELDIG
Helsinki Centre for Digital Humanities

HELSINKI INSTITUTE FOR
SOCIAL SCIENCES
AND HUMANITIES