

Aalto University, Espoo, Finland, Nov 28, 2024

# Welcome

## Tutorial: How to create a Linked Open Data service and semantic portal for your Cultural Heritage data

Eero Hyvönen

Aalto University, Dept. of Computer Science, professor  
University of Helsinki, Helsinki Centre for Digital Humanities (HELDIG), director

Semantic Computing Research Group (SeCo)

<http://seco.cs.aalto.fi/u/eahyvone/>





# Background: Linked Data work in the Finnish research infrastructure

**FIN-CLARIAH / DARIAH-FI:** <https://seco.cs.aalto.fi/projects/fin-clariah/>



Network for Data-Intensive Social Science and Humanities Research

 **DARIAH-EU**  
Celebrating the first decade

☰ 🔍

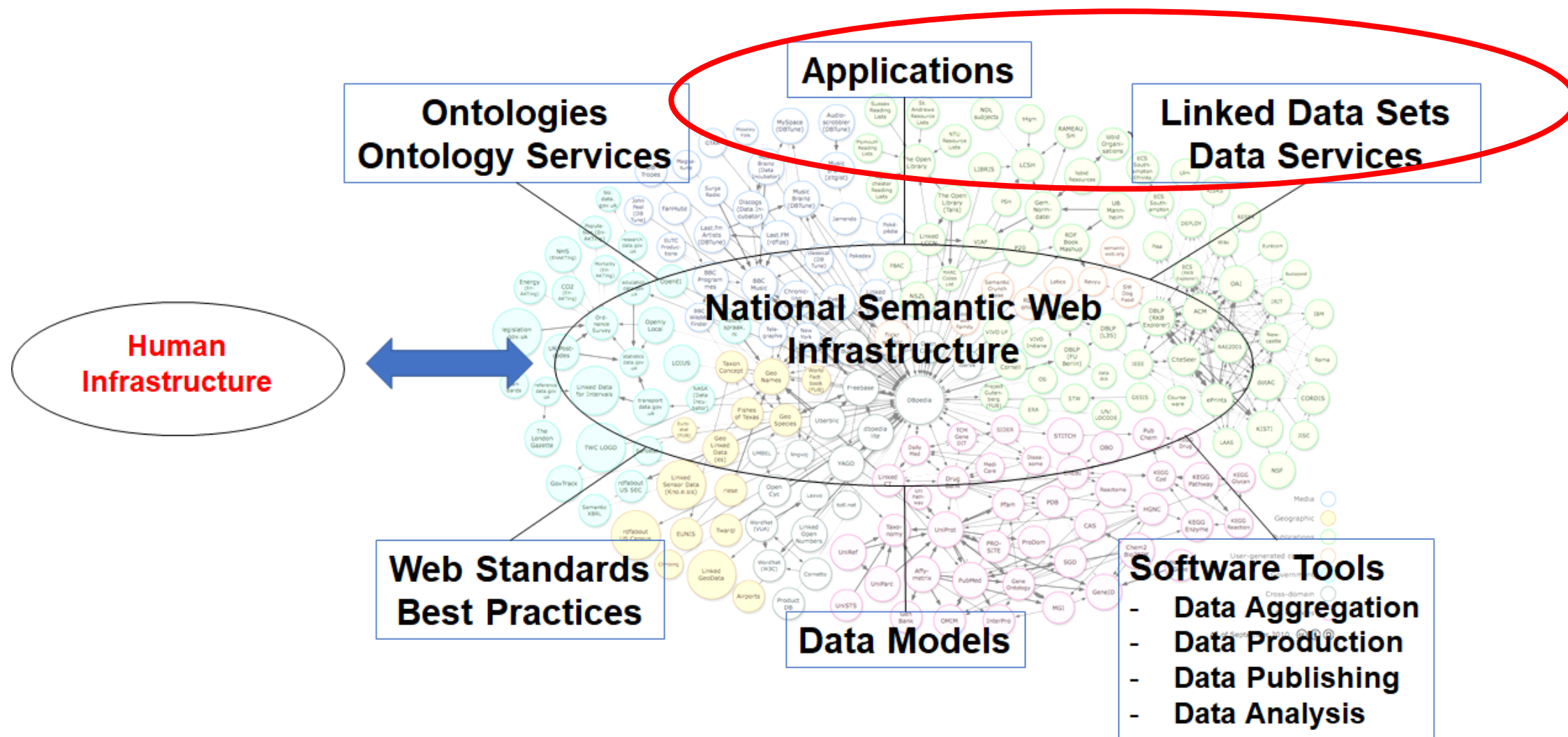
**DARIAH-FI**  
<https://www.dariah.fi/>

The pan-European infrastructure for arts & humanities scholars

[Learn More About DARIAH](#)



# Elements of a National Linked Data Infrastructure



# Learning objectives

Explain in practice how to

1. Create a Linked Open Data (LOD) service on the Web
2. Use the LOD service for Digital Humanities research
3. Use the LOD service for creating (portal) applications

**Approach: Using Sampo model, tools, and data**

Tanks to our funders:



Funded by the  
**European Union**  
NextGenerationEU



# Program

After each presentation, a short QA session

Feel free to use the chat channel!

## Introduction

- 13:00–13:15 **Why use Linked Open Data: Sampo model for creating LOD services and semantic portals (15min)**  
Eero Hyvönen

## Creating Linked Open Data and Service

- 13:15–13:30 **Metadata models and ontologies for knowledge representation (15min)**  
Jouni Tuominen
- 13:30–13:45 **Creating LOD from databases (15min)**  
Annastiina Ahola, Heikki Rantala
- 13:45–14:00 **Semantic enrichment: data linking, reasoning, and knowledge extraction from texts (15min)**  
Petri Leskinen, Rafael Leal
- 14:00-14:10 **Publishing Sampo data (dump), data service (LDF.fi), and portal on the Web (10min)**  
Jouni Tuominen
- 14:10-14:25 **Break (15min)**

## Using the Data Service and Portal

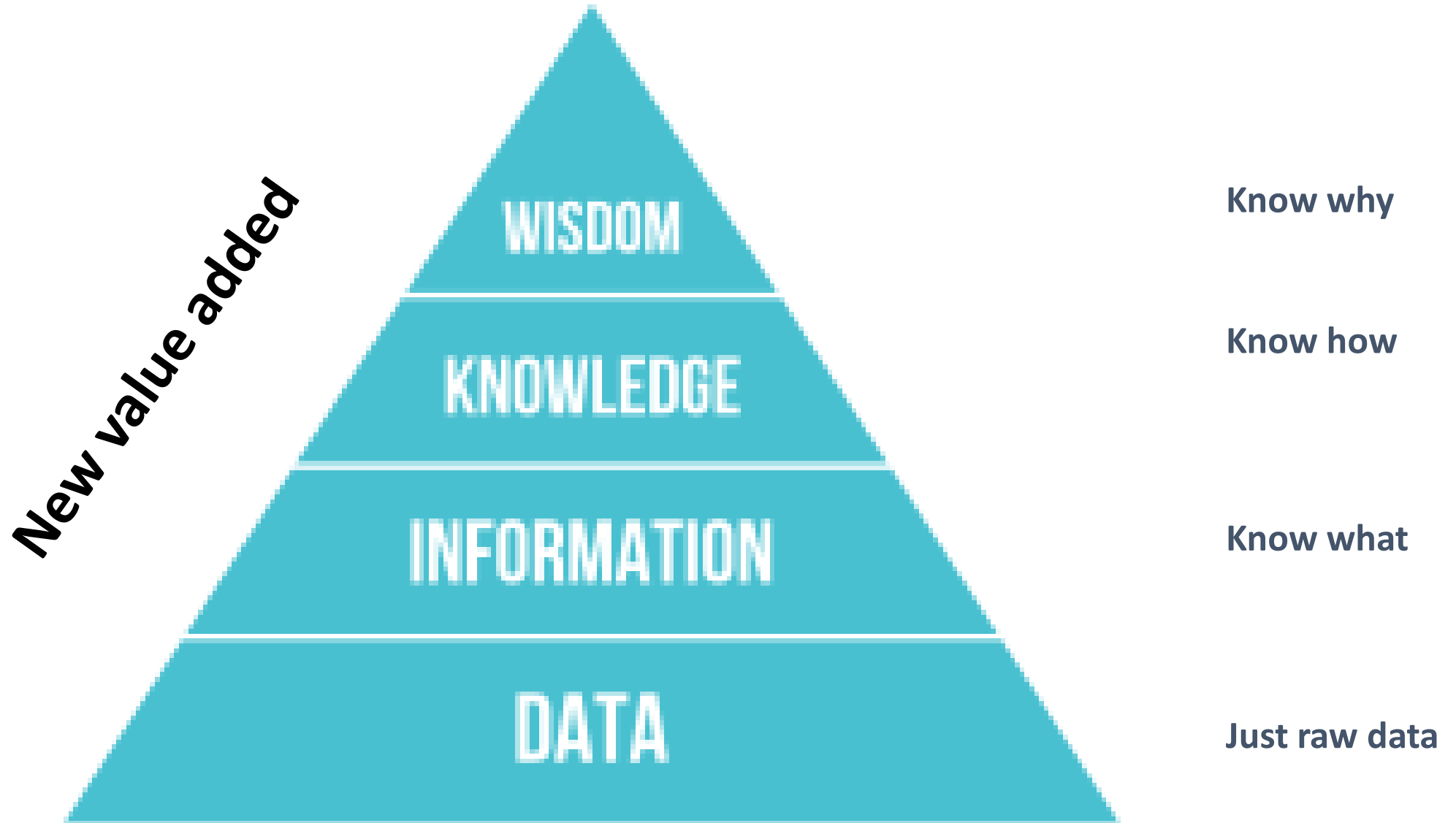
- 14:25-14:40 **Using the LOD service via SPARQL and Notebooks, data analysis, network analysis, and visualizations (15min)**  
Petri Leskinen
- 14:40-15:00 **Creating a semantic portal easily using Sampo-UI (20min)**  
Annastiina Ahola, Heikki Rantala
- 15:00-15:10 **Using the portal for Digital Humanities research, ParliamentSampo as an example (10min)**  
Heikki Rantala
- 15:10-16:00 **Networking with a glass of wine**

# **Why use Linked Open Data? Sampo model for creating LOD services and semantic portals**

**Eero Hyvönen**

# Semantic Web = Web of AI and Wisdom

Data -> Information -> Knowledge -> Wisdom (DIKW) hierarchy of Information Science



# Semantic Web is Based on Linked (Open) Data

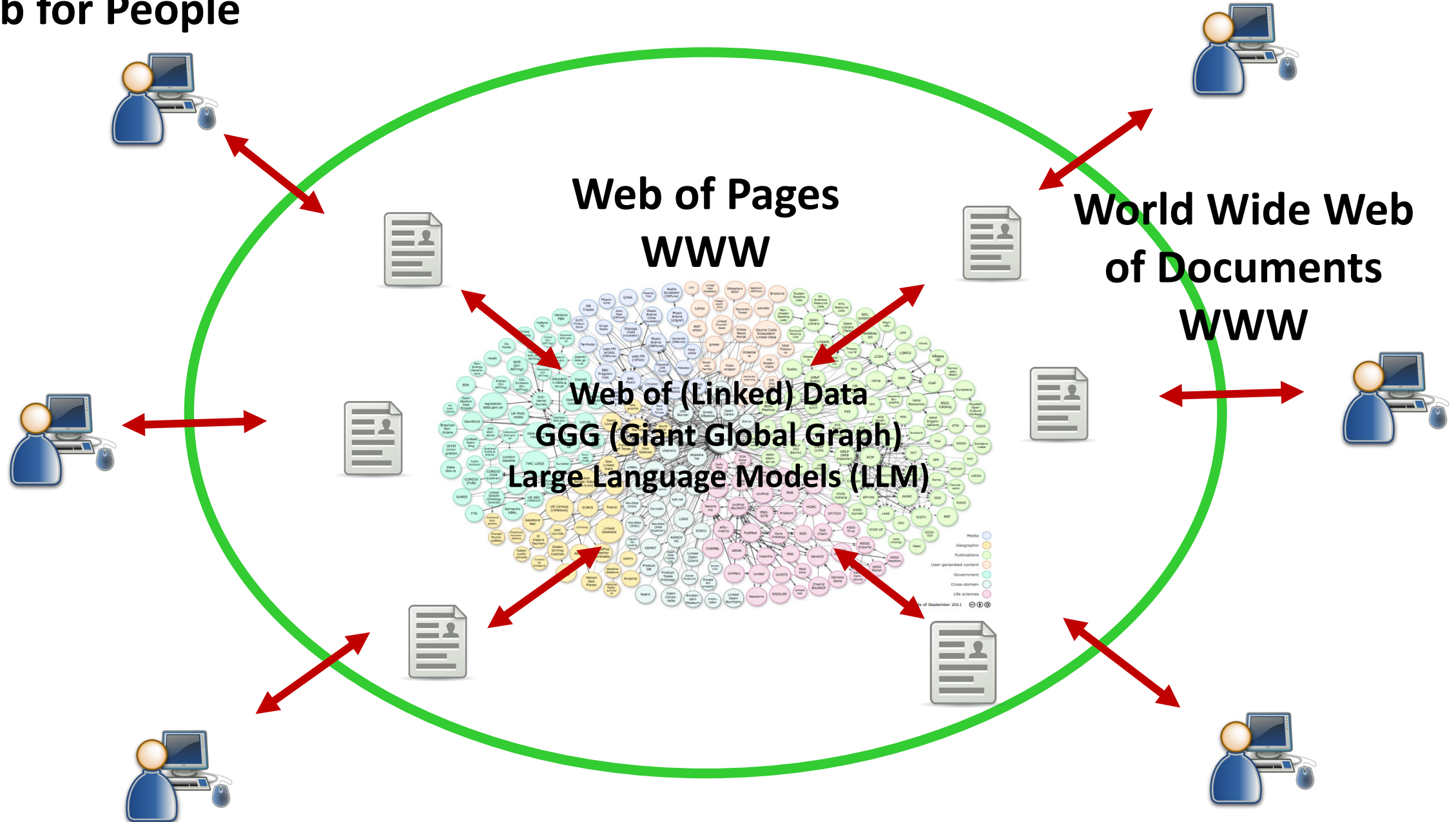
*“Learn how to see. Realize that everything connects to everything else.”*

Leonardo da Vinci (1452-1519)



# Semantic Web: WWW based on GGG/KGs

Web for People



# Why Linked (Open) Data

- Enriching everybody's data collaboratively from separate silos
  - Everybody wins by collaboration!
- Creating **F**indable, **A**ccessible, **I**nteroperable, **R**e-usable data
  - The value of data increases!
- Creating more intelligent applications for the public, curators, and researchers
  - The machine “understands” linked data!



**FAIR**

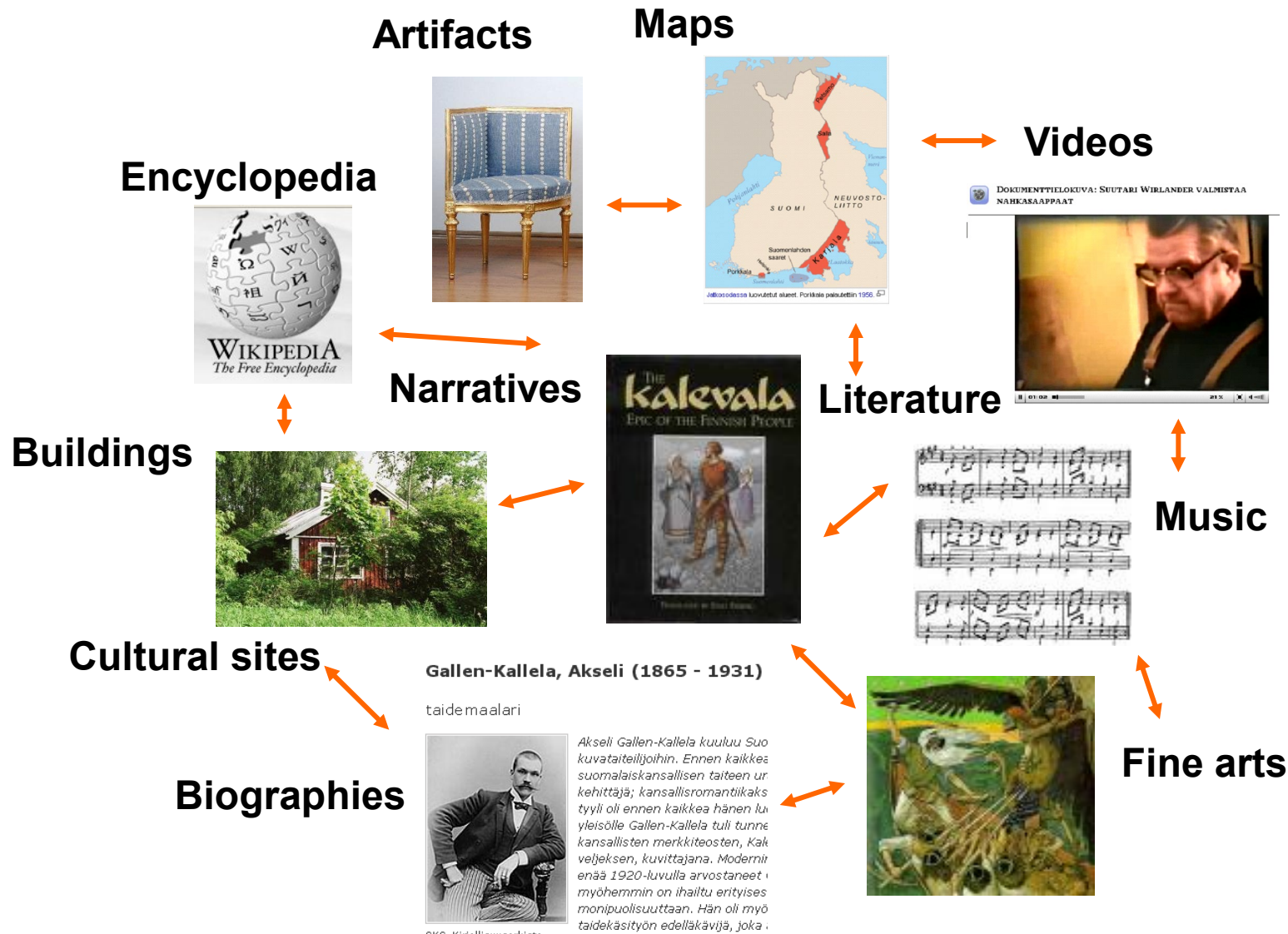


<https://www.go-fair.org/fair-principles/>



# Key Problem 1 in Cultural Heritage (CH) Content: CH Content is Semantically Complex and Linked

- **Heterogeneous contents**
  - text, speech, images, 3D, movies, animations,...
- **Incompatible data**
  - formats & models
- **Multi-lingual**
- **Semantically interlinked**





# Key Problem 2: CH Content Production System

- **Distributed silos**
- **Independent actors**
- **Little collaboration**

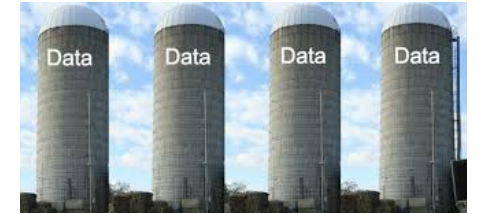
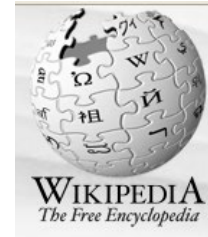
Land survey



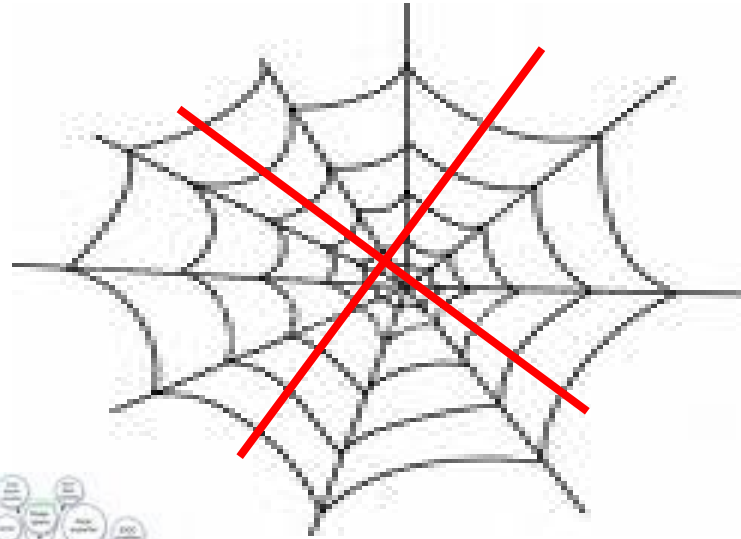
Museums



Web 2.0 sites



Archives



Media



Linked Data



Citizens

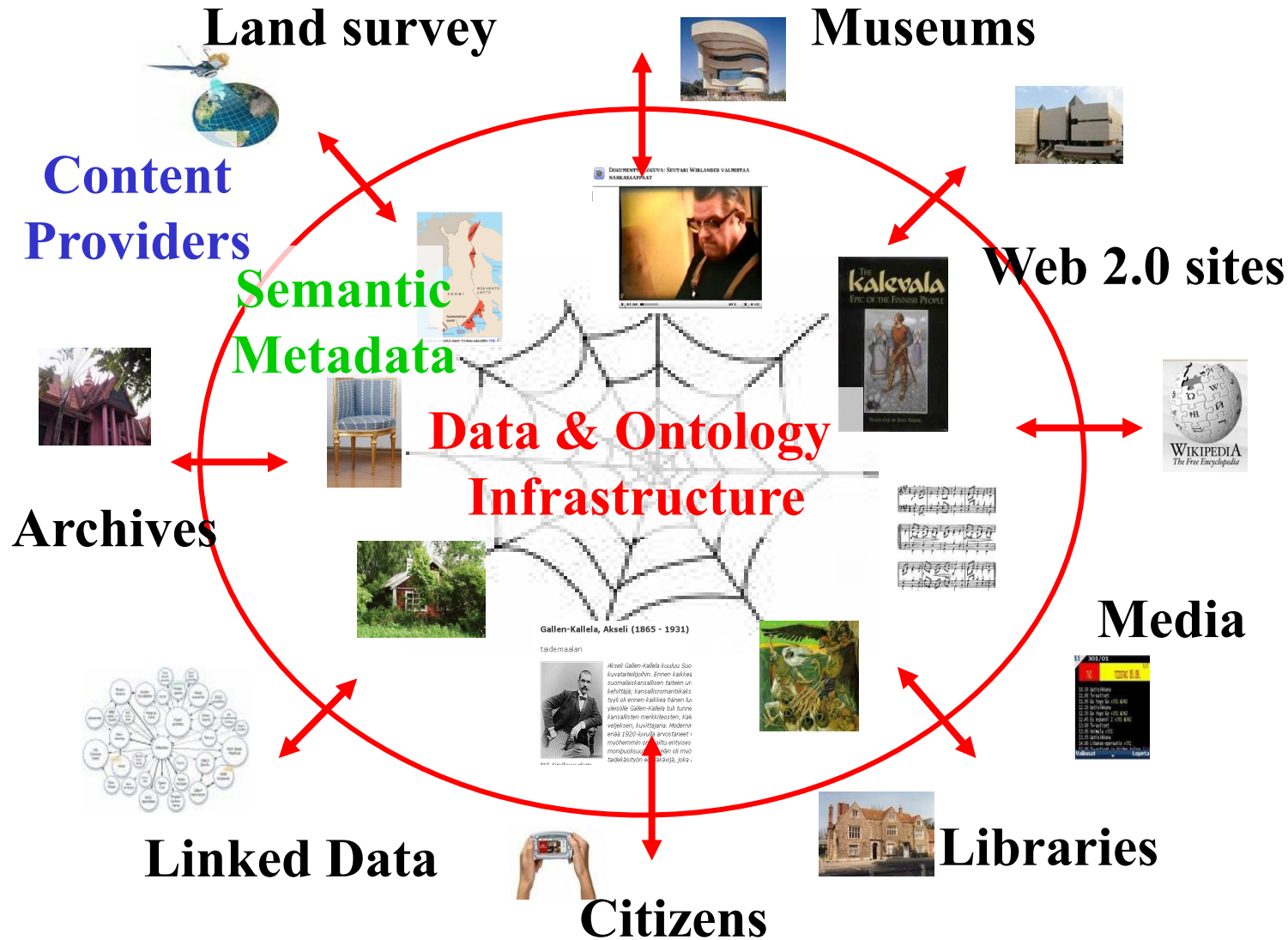


Libraries





# Solution: Collaborative Cultural Heritage Publishing Model “Sampo”



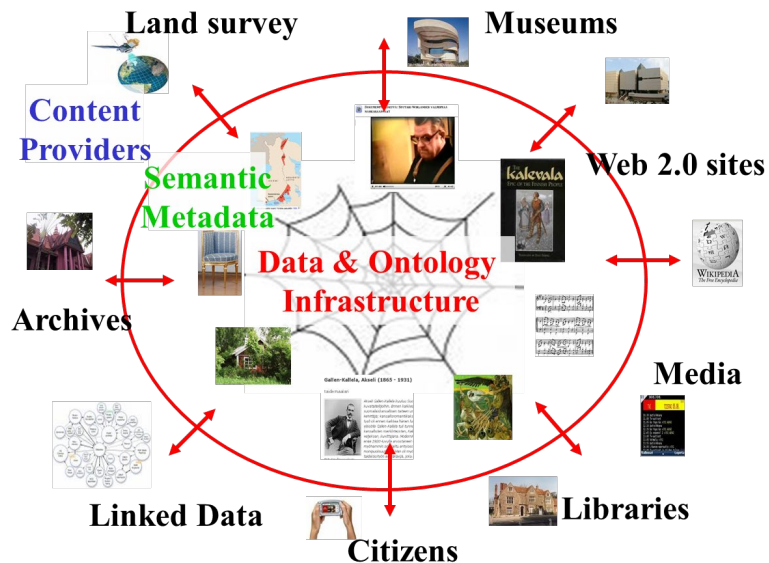
Findable  
Accessible  
Inteoperable  
Re-usable data!

# Sampo Model for 1) Human Users and 2) Machines



## LINKED DATA SERVICE

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



Linked Data Finland

Living Laboratory Data Service for the Semantic Web

<https://ldf.fi/>

## APPLICATIONS & TOOLS

Use Cases

Tools

### Sampo Portals & Applications

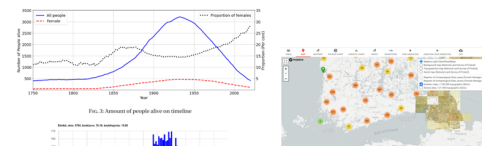


Sampos:

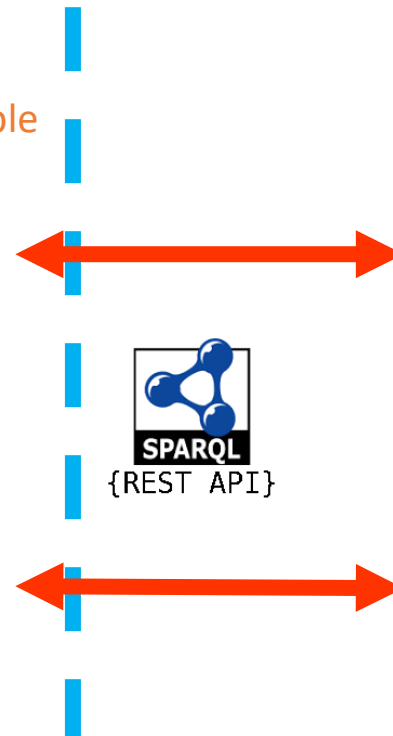
<https://seco.cs.aalto.fi/applications/sampo/>

Sampo-UI, etc.

### Digital Humanities Research



Google Colab, Jupyter,  
YASGUI, etc.



### 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.

# Sampo Model Principles

## Sampo Model Principles P1–P6



### Principles for creating LOD services

How to create a  
KG & data service

### Principles for creating portal interfaces

How to create a  
KG –based portal UI

P1	Support collaborative data creation and publishing
P2	Use a shared open ontology infrastructure
P3	Make clear distinction between the LOD service and the user interface (UI)
P4	Provide multiple perspectives to the same data
P5	Standardize portal usage by a simple filter-analyze two-step cycle
P6	Support data analysis and knowledge discovery in addition to data exploration

[Hyvönen, *Semantic Web*, 2023]  
[Ikkala et al., *Semantic Web*, 2022]



Eero Hyvönen: [Digital Humanities on the Semantic Web: Sampo Model and Portal Series](#).  
*Semantic Web – Interoperability, Usability, Applicability*, vol. 14, no. 4, pp. 729-744, 2023.

# Why infrastructure and FAIR data?

“Intellectuals solve problems  
- geniuses prevent them”

Albert Einstein

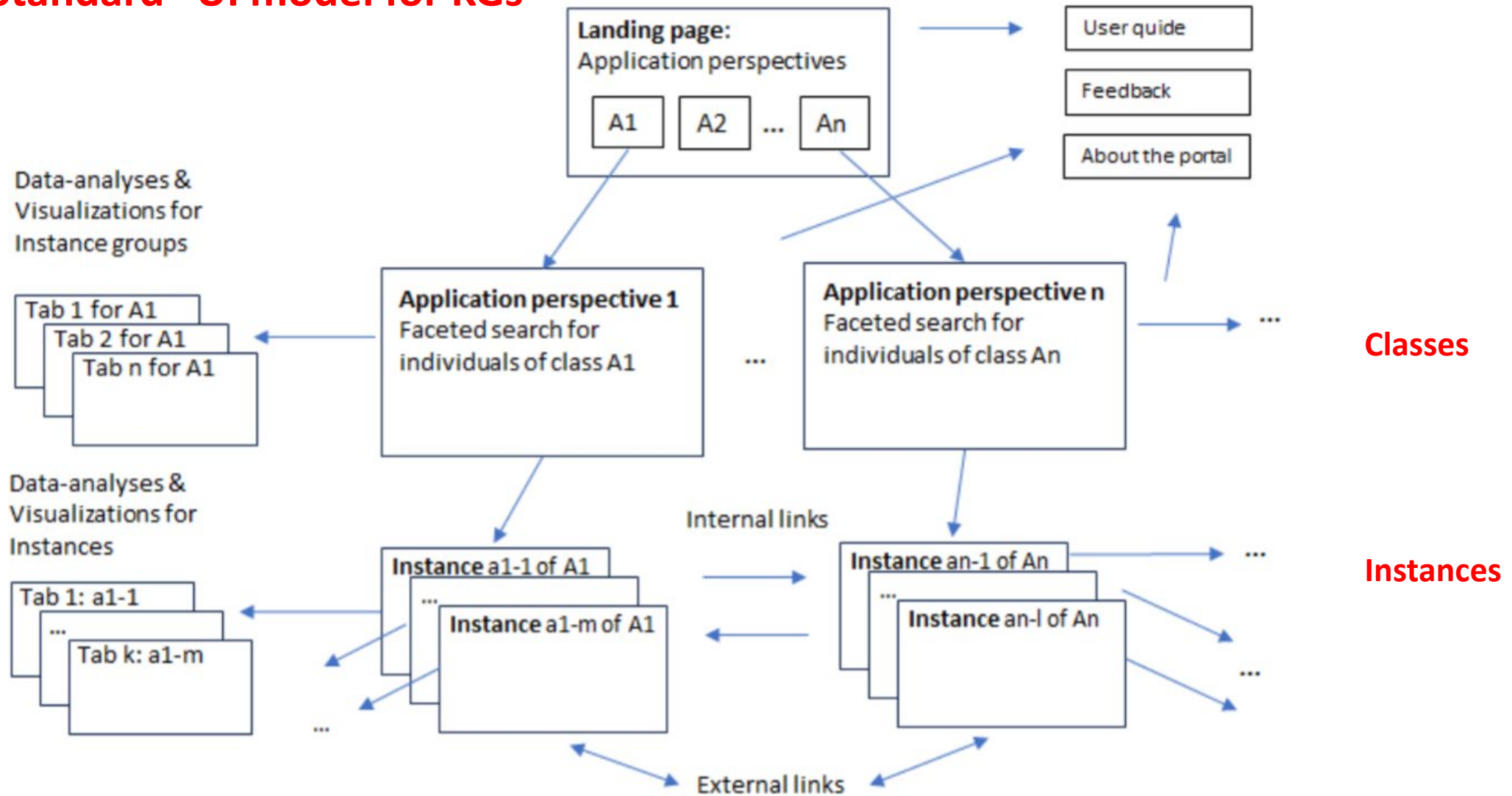
FAIR Data



# Sampo UI: Full Stack JavaScript framework for developing semantic portals

- **Key ideas:**
  - All you need is a SPARQL endpoint! Data service detached from UI
  - Integrate faceted search and browsing seamlessly with data analyses & visualizations
- **Easy to use**
  - For developers
  - For end users
- **More info**
  - 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, pp. 69-84, 2022.
  - Heikki Rantala, Annastiina Ahola, Esko Ikkala and Eero Hyvönen: [How to create easily a data analytic semantic portal on top of a SPARQL endpoint: introducing the configurable Sampo-UI framework](#). *VOILA! 2023 Visualization and Interaction for Ontologies, Linked Data and Knowledge Graphs 2023*, CEUR WS Proceedings, Vol. 3508, 2023.
  - GitHub: <https://github.com/SemanticComputing/sampo-ui>
  - Tutorials etc: <https://seco.cs.aalto.fi/tools/sampo-ui/>

# Sampo-UI “Standard” UI model for KGs



**Figure 1:** Navigational page structure of a portal based on Sampo-UI.

# Sampo Series of LOD Services & Portals



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

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)

Work in progress: **ArtSampo**, **LetterSampo Finland (1809-1917)**, **PASampo**, **SampoSampo**, ...



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

# "Sampo Series" Demonstrates a Paradigm Shift: 4 Generations of Publishing Data for Humanities

1. Carvings & 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



# More Information – Questions?

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

550 publications about the infras and their applications:

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

Videos and applications:

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



Semantic Computing Research Group (SeCo)

Making computers and the web more intelligent and interoperable

## HOMEPAGE

MISSION

NEWS

PROJECTS

EVENTS

PEOPLE

PARTNERS

CAREERS

TEACHING

CONTACT

PUBLICATIONS

APPLICATIONS

ONTOLOGIES

LINKED DATA

INFRA FOR DH

SERVICES

TOOLS

SOFTWARE

DOWNLOAD

Wikipedia: *Semantic web suomi*

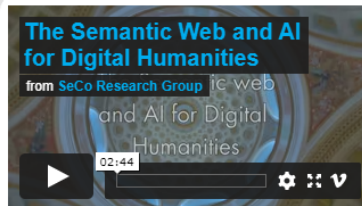
Google this site

## SeCo Welcomes You!

Semantic Computing Research Group (SeCo) researches machine-processable semantics related to, e.g., the Semantic Web. We are located at the [Aalto University, School of Science, Department of Computer Science](#) and [HELDIG – Helsinki Centre for Digital Humanities](#), University of Helsinki, Faculty of Arts. Our group consists of researchers both from the Aalto University, School of Science, Department of Computer Science and [University of Helsinki, Department of Computer Science](#), and HELDIG.

Our research is [focused](#) on semantic technologies, such as the Semantic Web and intelligent web services. In addition to research and [publications](#), we also create prototype applications that demonstrate the new possibilities of semantic technologies, such as **semantic portals for end users**, **semantic infrastructural data and ontology services**, and **ontologies and tools** for creating semantic applications.



## Application Domains

Our work is highly cross-disciplinary, including application domains, such as Digital Humanities, Health, Learning, Government, Commerce, Geography, and Biology.

## Selected SeCo Applications and Demos for End Users



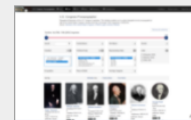
[Mapping Manuscript Migrations](#)  
Online since 2020. [Try it!](#)



[WarVictimSampo 1914-1922](#)  
Online since 2019. [Try it!](#)



[BiographySampo - Finnish Life Stories on the Semantic Web](#)  
Online since 2018. [Try it!](#)



[U.S. Congress Prosopographer](#)  
Online since 2018. [Try it!](#)



[Norssi Alumni on the Semantic Web](#)  
Online since 2017. [Try it!](#)



[Semantic Finlex - Finnish Law](#)



[WarSampo - Second World War](#)



[Hipla.fi - Finnish Ontology](#)



[Semantic National Biography](#)



[BirdWatch - Mobile Semantic](#)

## Latest News

[Five new SeCo papers accepted for the ISWC 2021](#)  
The 20th International Semantic Web Conference (ISWC 2021), the...

2.8.2021 6:53 by eahyvone

[FindSampo ? Finnish Archaeological Finds on the Semantic Web](#)

FindSampo – Finnish Archaeological Finds on the Semantic Web, a...

19.5.2021 20:00 by eahyvone

[In/Tangible European Heritage: New EU project starting in SeCo](#)

The idea of the new 3-year project is to study how intangible cultural...

6.8.2020 8:57 by eahyvone

## Latest Publications

Rafael Leal, Joonas Kesäniemi, Mikko Koho and Eero Hyvönen: [Relevance Feedback Search Based on Automatic Annotation and Classification of Texts](#)

Eero Hyvönen: [How to Create a National Cross-domain Ontology and Linked Data Infrastructure and Use It on the Semantic Web](#)

Eero Hyvönen, Esko Ikkala, Mikko Koho, Jouni Tuominen, Toby Burrows, Lynn Ransom, and Hanno Wijsman: [Mapping Manuscript Migrations on the Semantic Web: A Semantic Portal and Linked Open Data Service for Premodern Manuscript Research](#)

Petri Leskinen and Eero Hyvönen: [Reconciling and Using Historical Person Registers as Linked Open Data in the AcademySampo Knowledge Graph](#)

## SeCo on Twitter

Tweets by @secoresearch

SeCo Research Group Retweeted



**ManuscriptMigrations**  
@MSMigrations

"Harmonizing and publishing heterogeneous premodern manuscript metadata as Linked Open