Digital Humanities in the Nordic and Baltic Countries, Tartu, Estonia, March 4, 2025

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

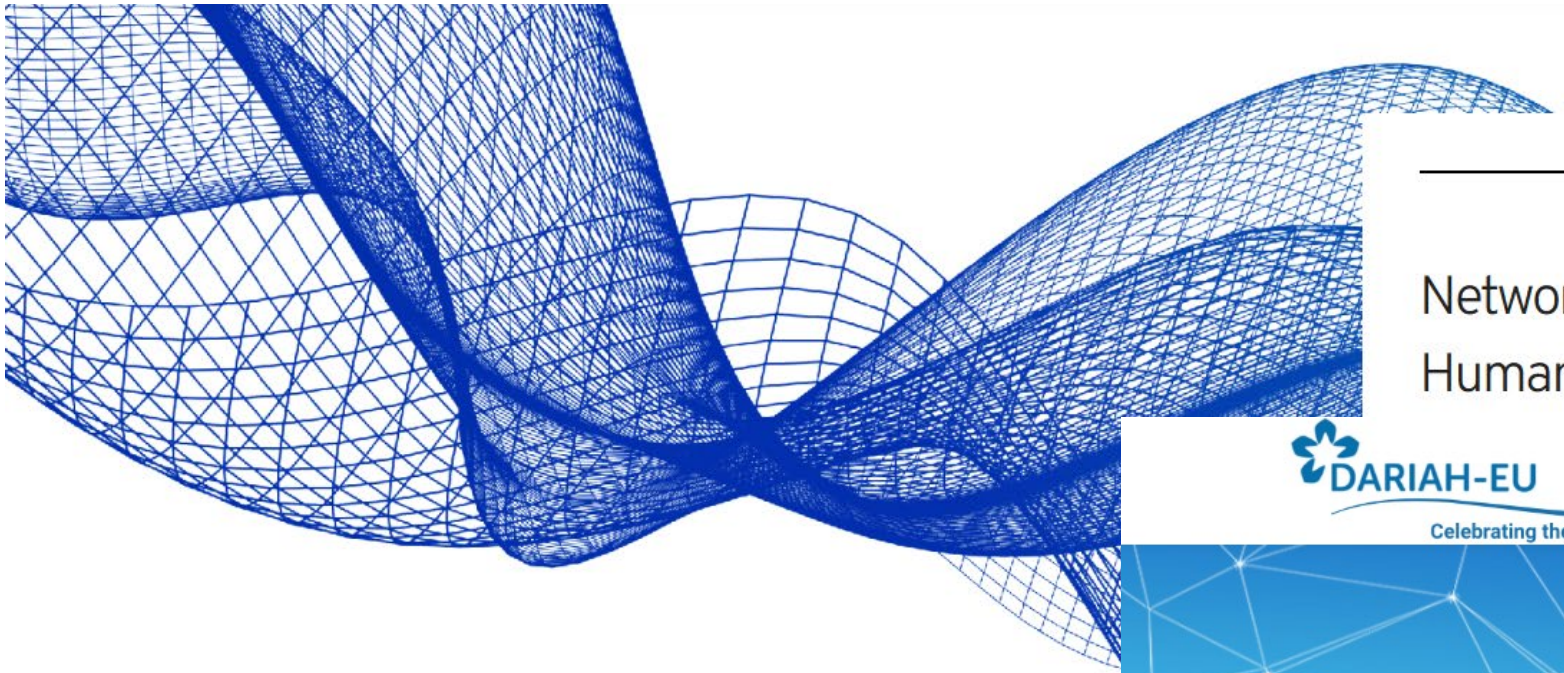
Semantic Computing Research Group (SeCo)

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/



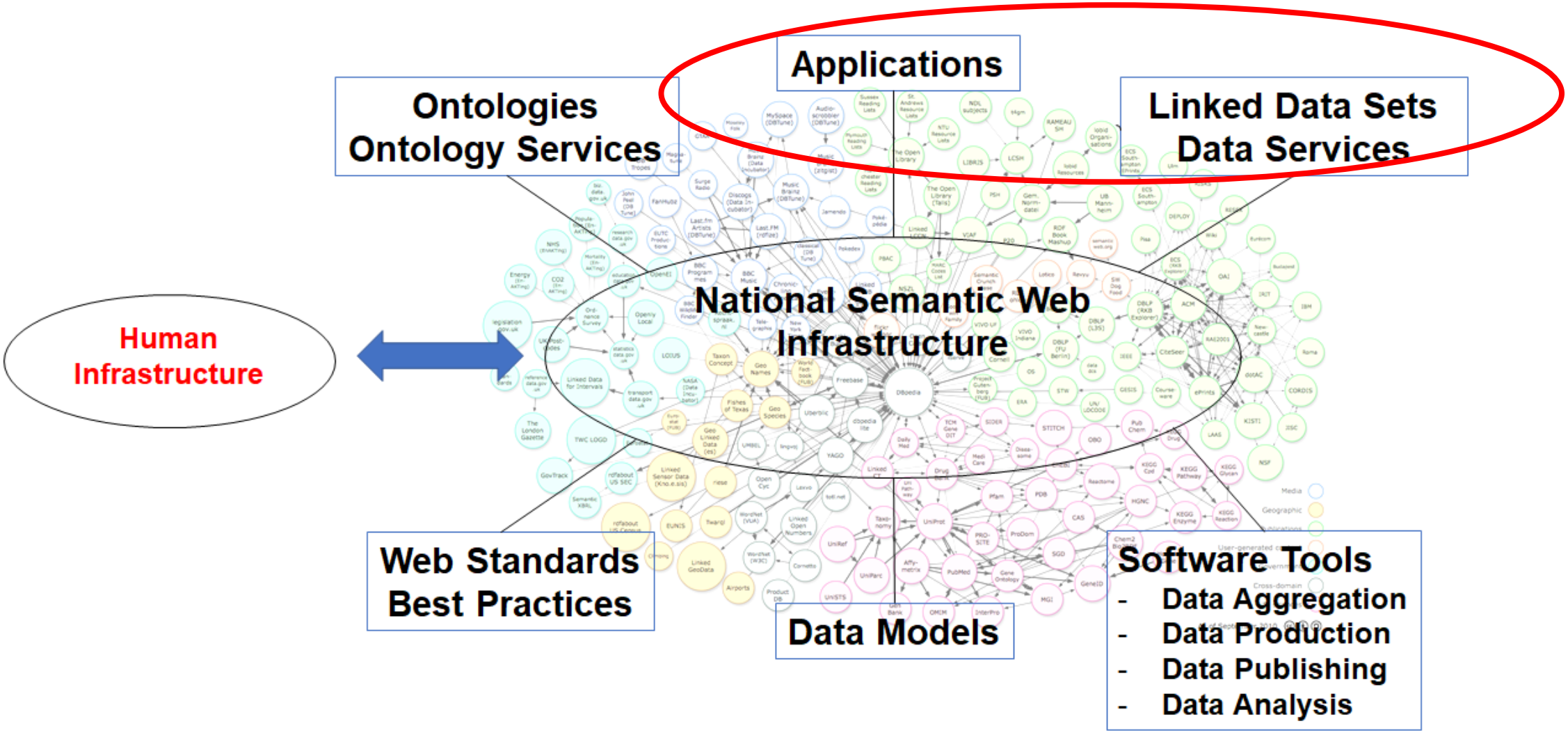
Network for Data-Intensive Social Science and Humanities Research

DARIAH-FI

https://www.dariah.fi/



Elements of a National Linked Data Infrastructure



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. bib pdf link

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:

Suomen Akatemia
Finlands Akademi
Research Council of Finland

Funded by the
European Union
NextGenerationEU

Program

Introduction

- 14:00–14:30 **Why use Linked Open Data: Sampo model for creating LOD services and semantic portals**
  Eero Hyvönen
- Example of a Sampo project and portal demo: LetterSampo - The Republic of Letters on the Semantic Web

Creating Linked Open Data and Service

- 14:30–15:45 **Metadata models and ontologies for knowledge representation**
  Jouni Tuominen
- 15:45–15:15 **Creating LOD from databases**
  Annastiiina Ahola, Heikki Rantala
- 15:15–15:45 **Semantic enrichment: data linking, reasoning, and knowledge extraction from texts**
  Petri Leskinen and Rafael Leal
- 15:45-16:00 **Publishing Sampo data (dump), data service (LDF.fi), and portal on the Web**
  Jouni Tuominen
- 16:00-16:30 **Coffee Break**

Using the Data Service and Portal

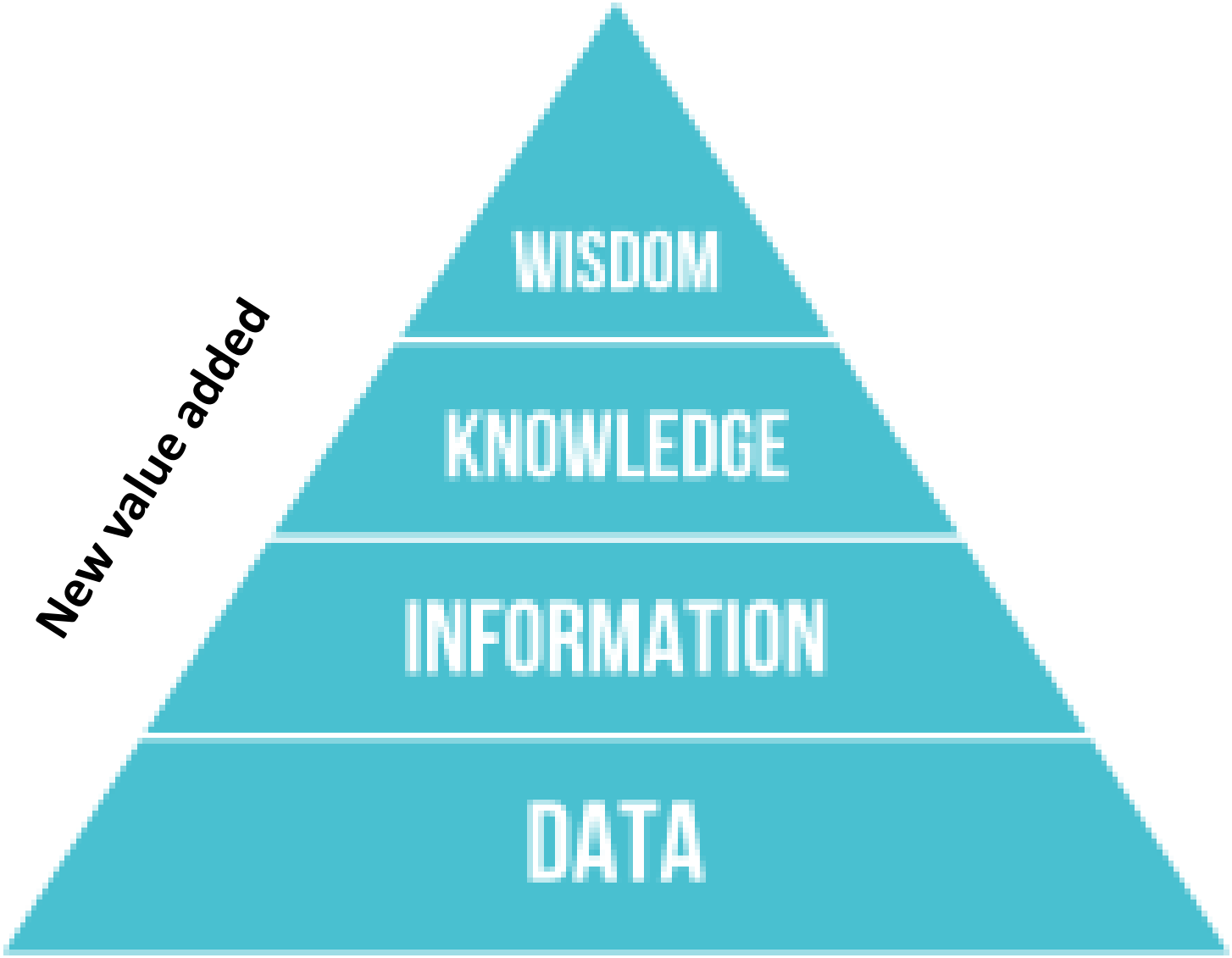
- 16:30-17:00 **Using the LOD service via SPARQL and Notebooks, data analysis, network analysis, and visualizations**
  Petri Leskinen
- 17:00-17:30 **Creating a semantic portal easily using Sampo-UI**
  Annastiina Ahola, Heikki Rantala
- 17:30-18:00 **Using the portal for Digital Humanities research: ParliamentSampo and other examples**
  Heikki Rantala

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



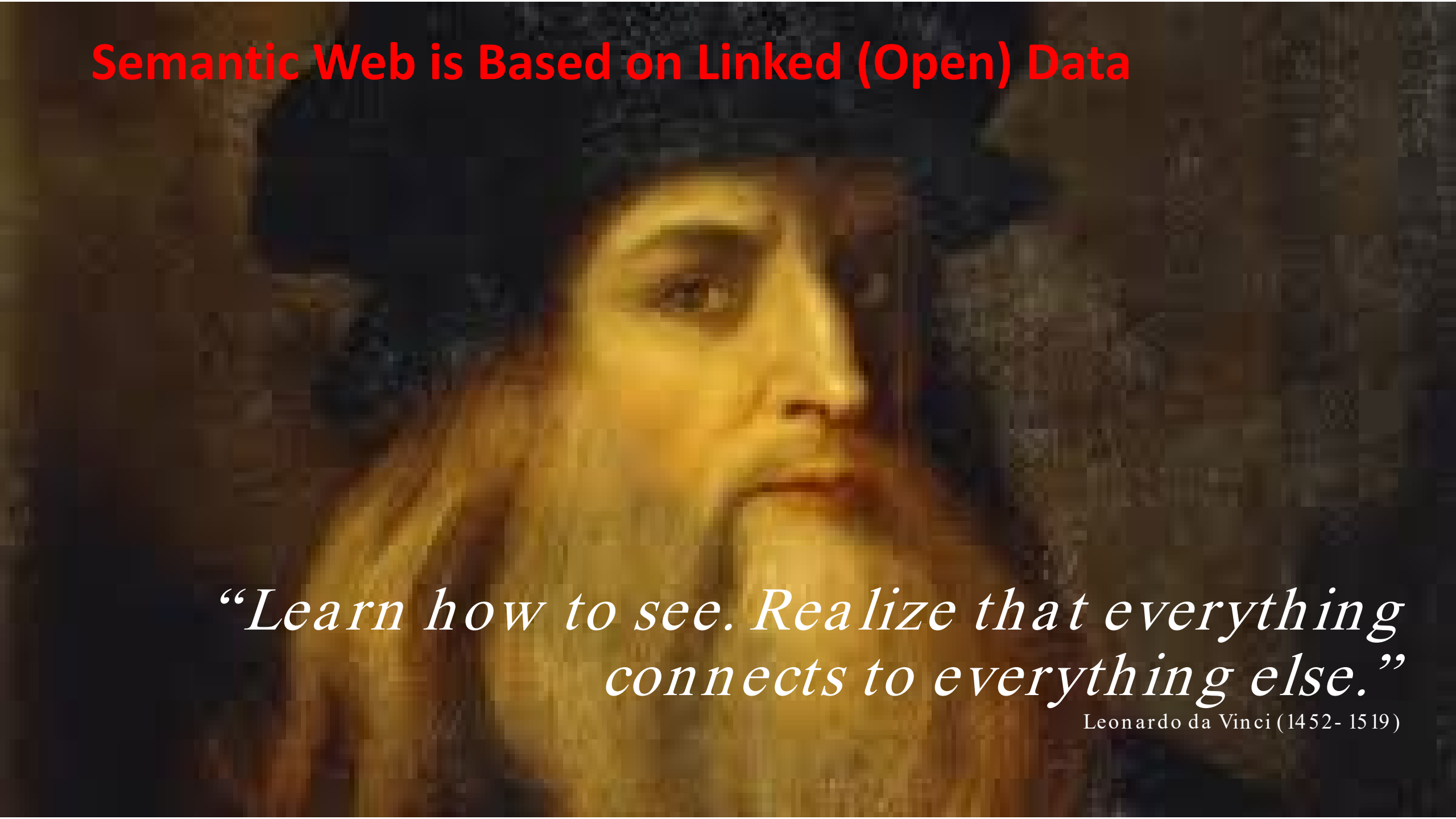
Know why

Know how

Know what

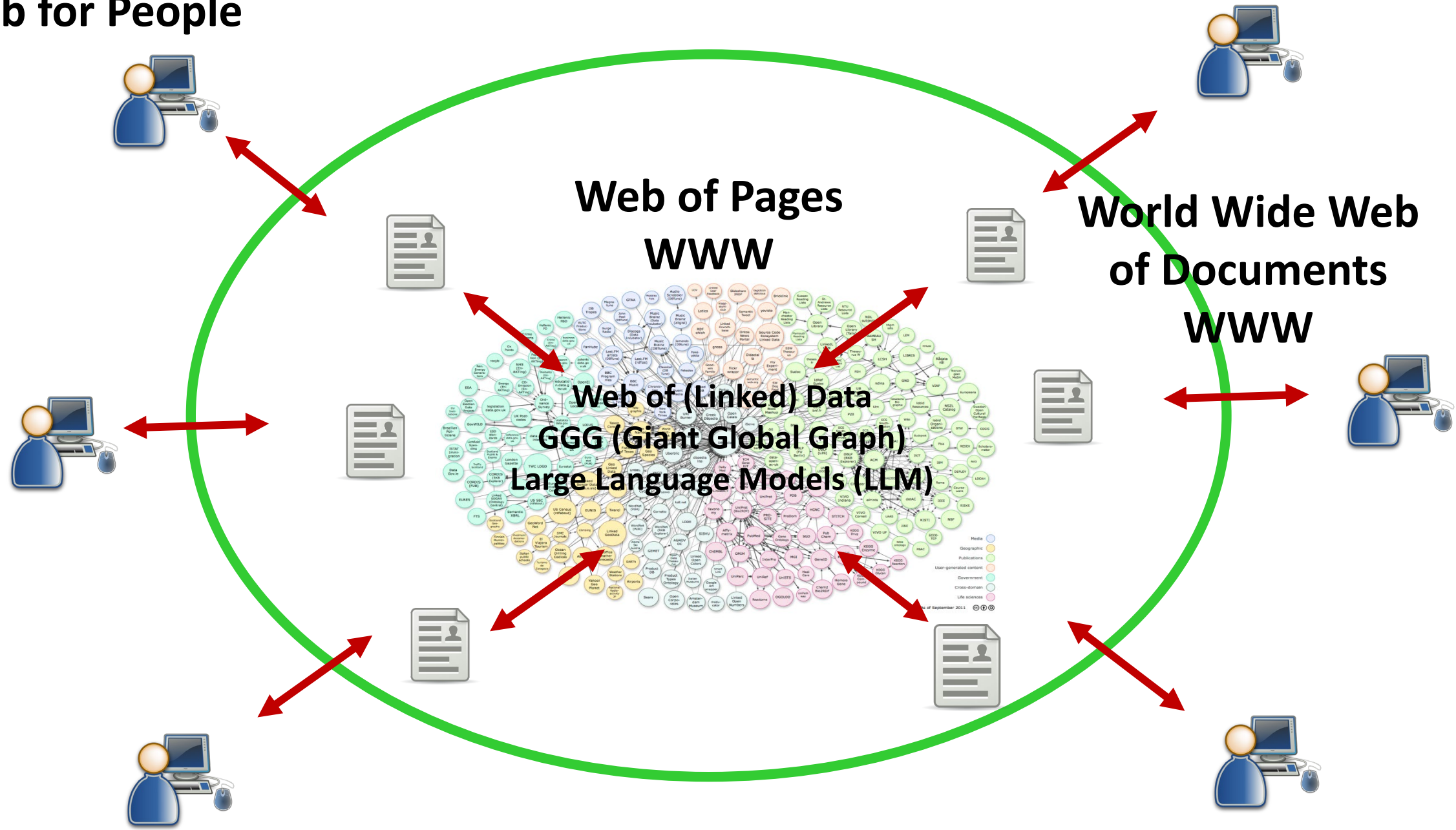
Just raw data

Semantic Web is Based on Linked (Open) Data



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 Findable, Accessible, Interoperable, Re-usable data
  - The value of data increases!

FAIR

GO FAIR

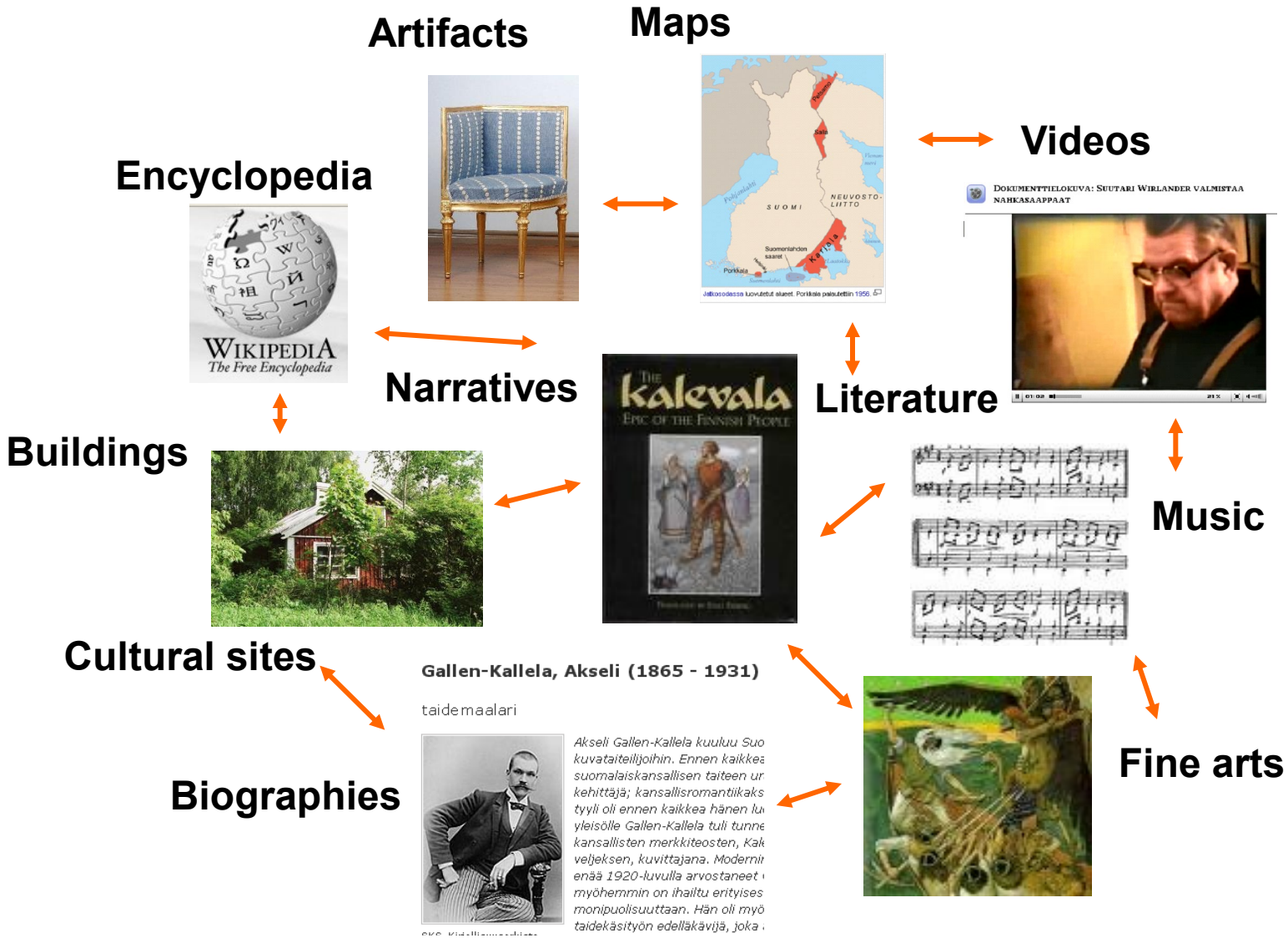
https://www.go-fair.org/fair-principles/

- Creating more intelligent applications for the public, curators, and researchers
  - The machine "understands" linked data!



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

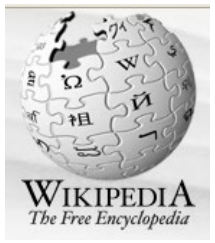


Museums





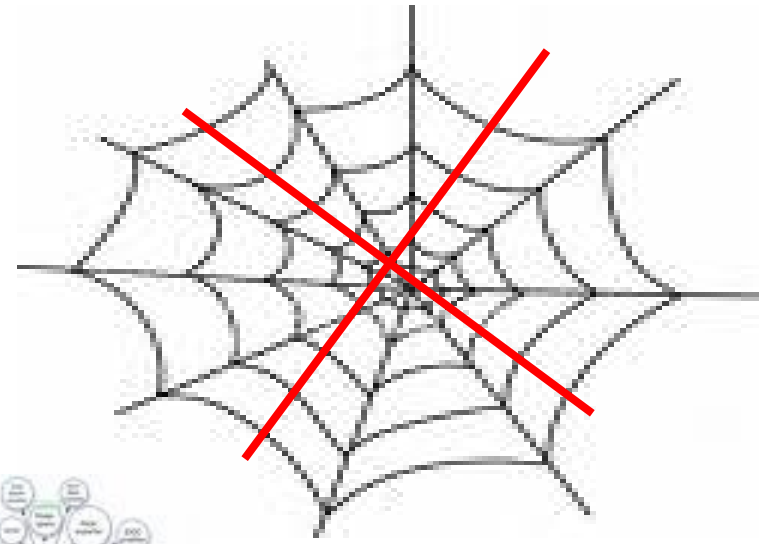
Web 2.0 sites





Archives





Media





Linked Data



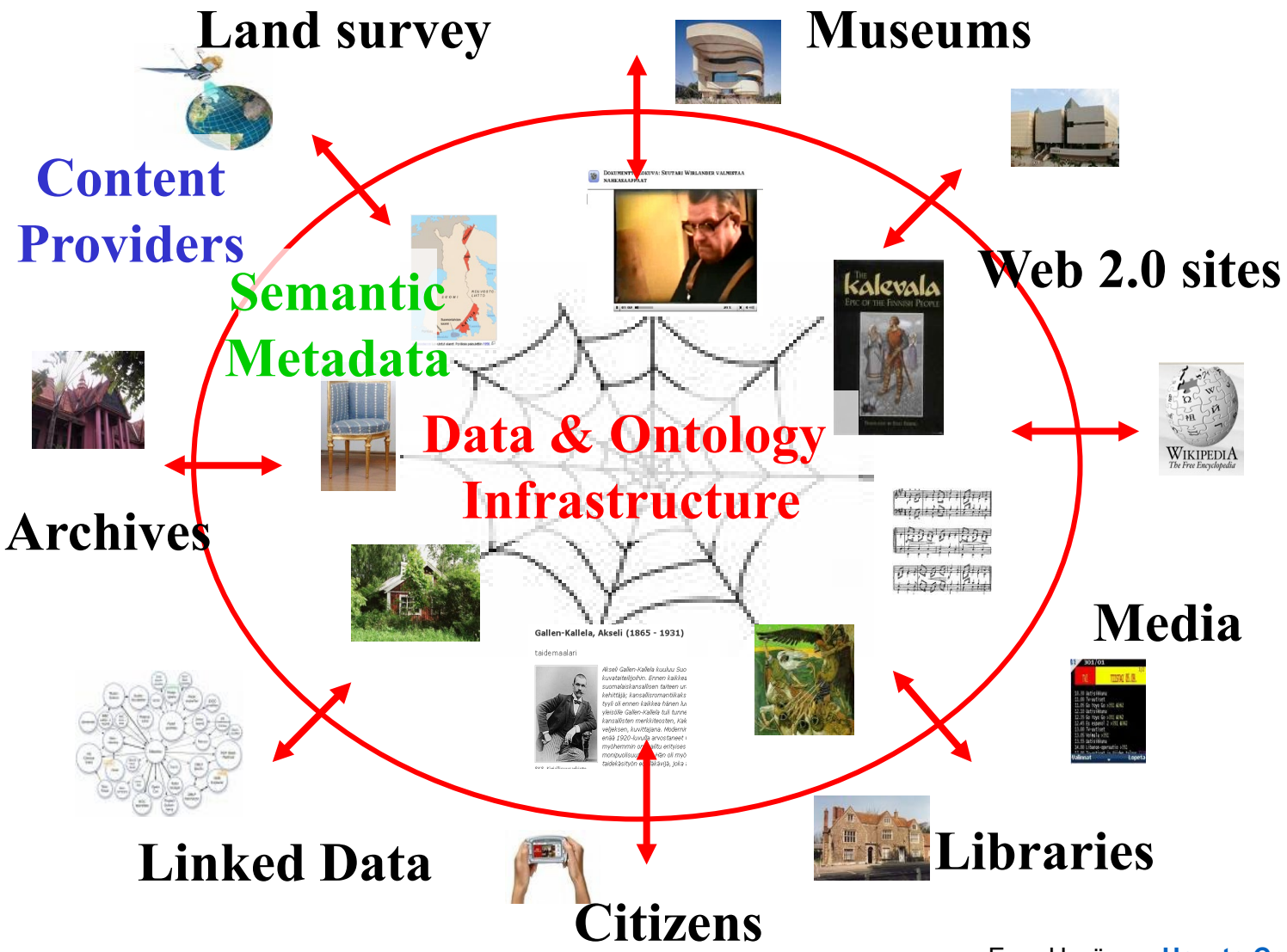


Citizens



Libraries

Solution: Collaborative Cultural Heritage Publishing Model “Sampo”



GO FAIR

Findable
Accessible
Inteoperable
Re-usable data!

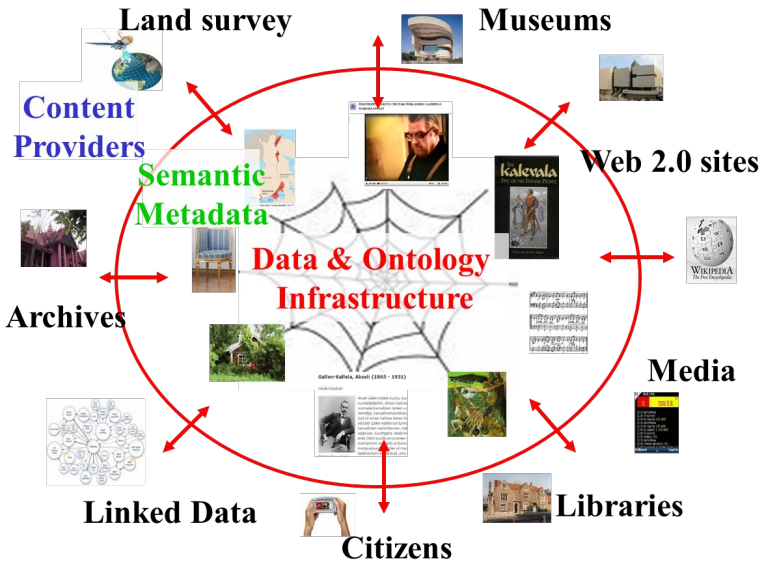
Eero Hyvönen: [How to Create a National Cross-domain Ontology and Linked Data Infrastructure and Use It on the Semantic Web](), Keynote at DCMI 2021 conference.

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/

SPARQL
{REST API}

APPLICATIONS & TOOLS

Use Cases

Tools

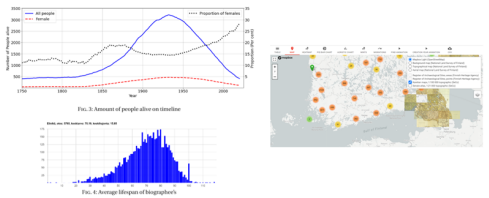
Sampo Portals & Applications



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

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



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

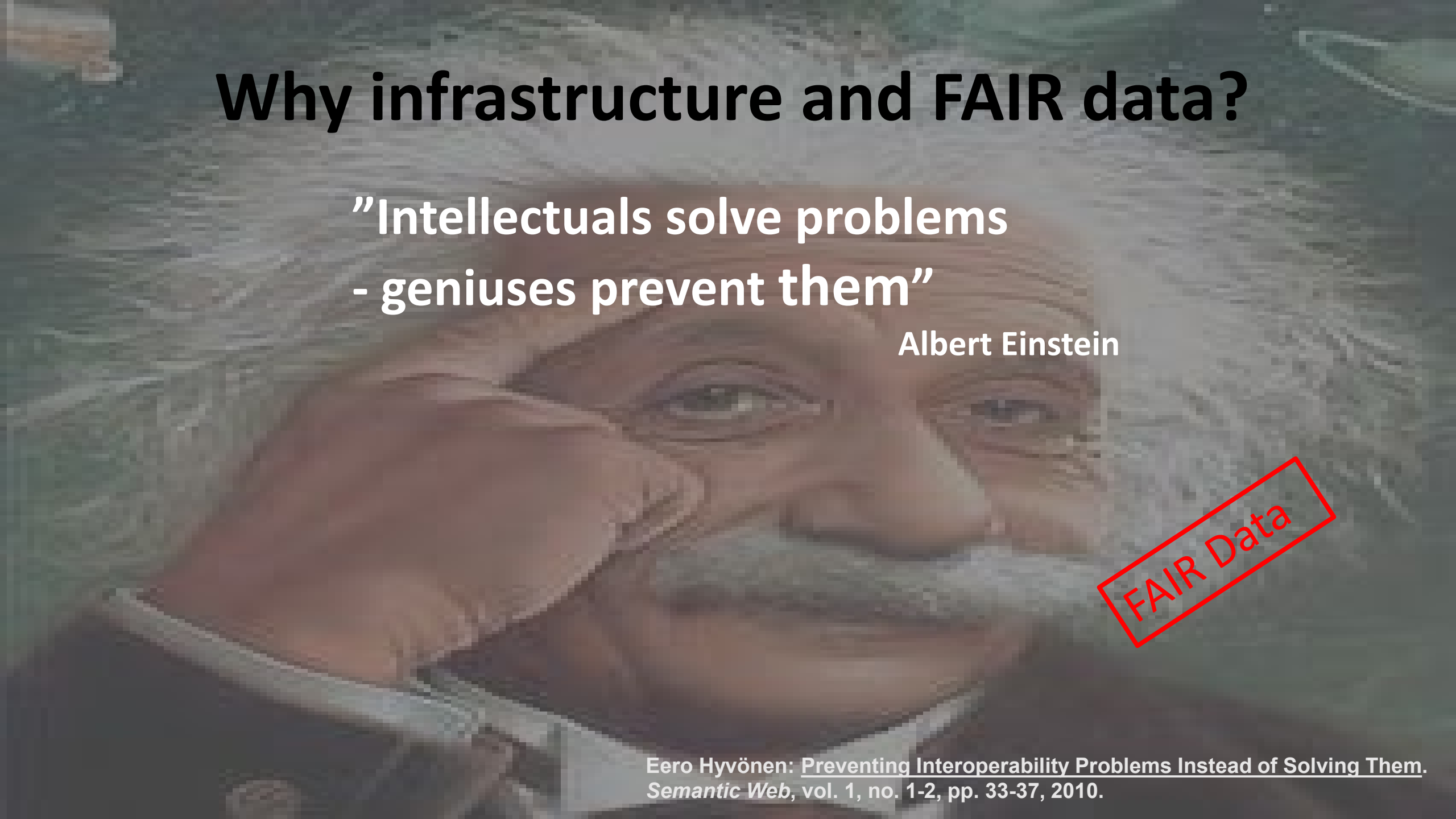
[Hyvönen, *Semantic Web*, 2023]
[Ikkala et al., *Semantic Web*, 2022]

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?

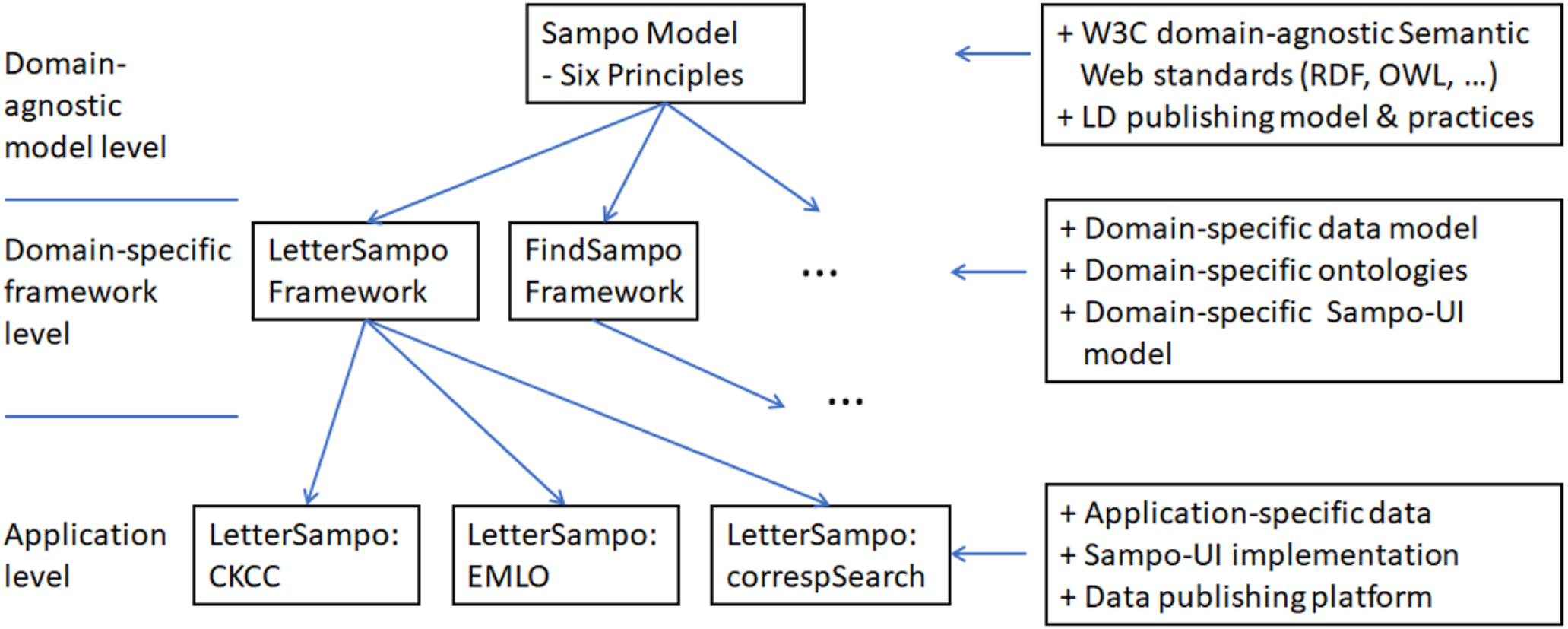
”Intellectuals solve problems
- geniuses prevent them”

Albert Einstein



Eero Hyvönen: Preventing Interoperability Problems Instead of Solving Them. *Semantic Web*, vol. 1, no. 1-2, pp. 33-37, 2010.

Sampo Frameworks: Adding domain specific level to Sampo Model



Eero Hyvönen, Petri Leskinen and Jouni Tuominen: **LetterSampo – Historical Letters on the Semantic Web: A Framework and Its Application to Publishing and Using Epistolary Data of the Republic of Letters**. Journal on Computing and Cultural Heritage, vol. 16, no. 1, 2023. bib pdf link

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

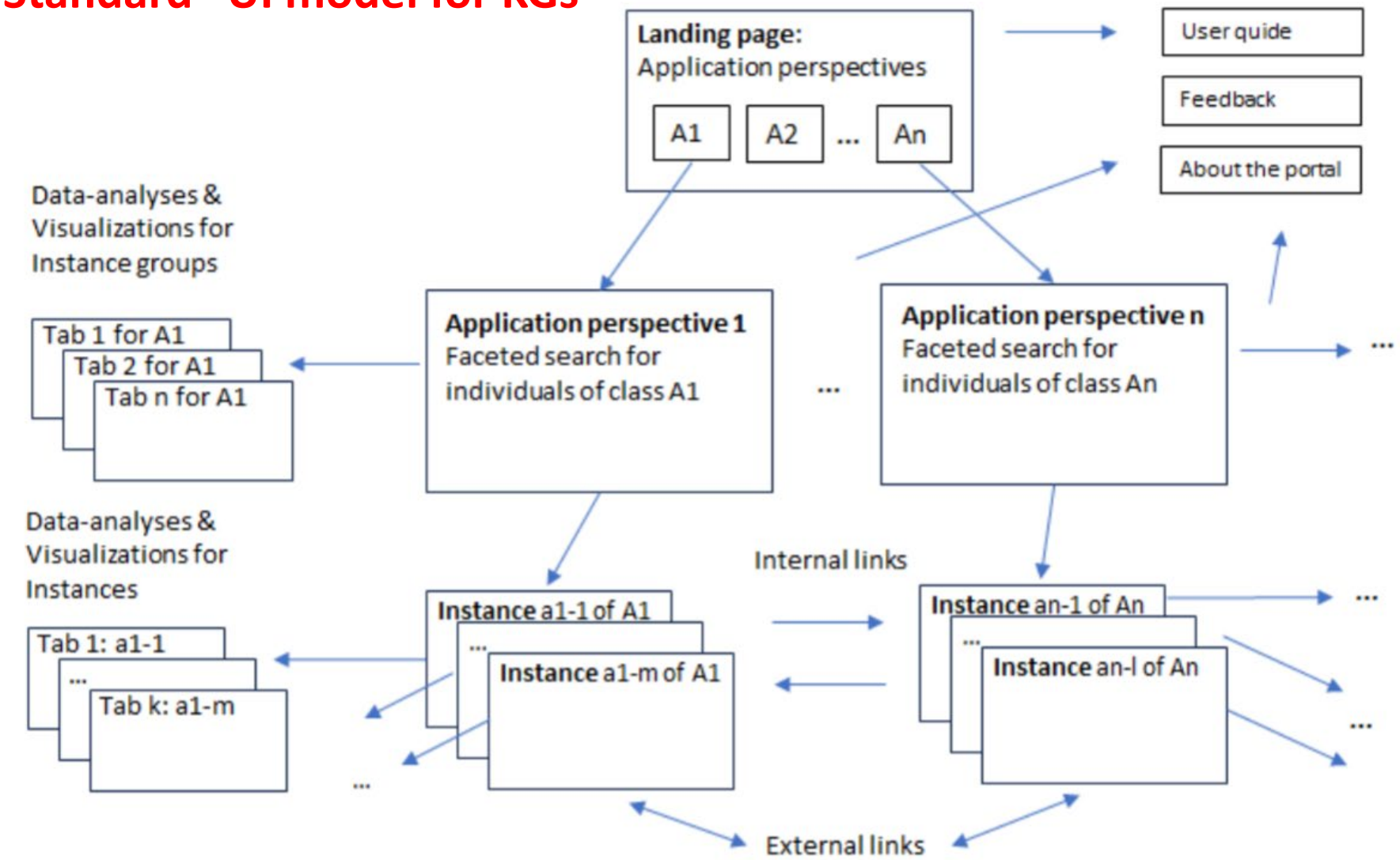


Figure 1: Navigational page structure of a portal based on Sampo-UI.

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.***

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

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



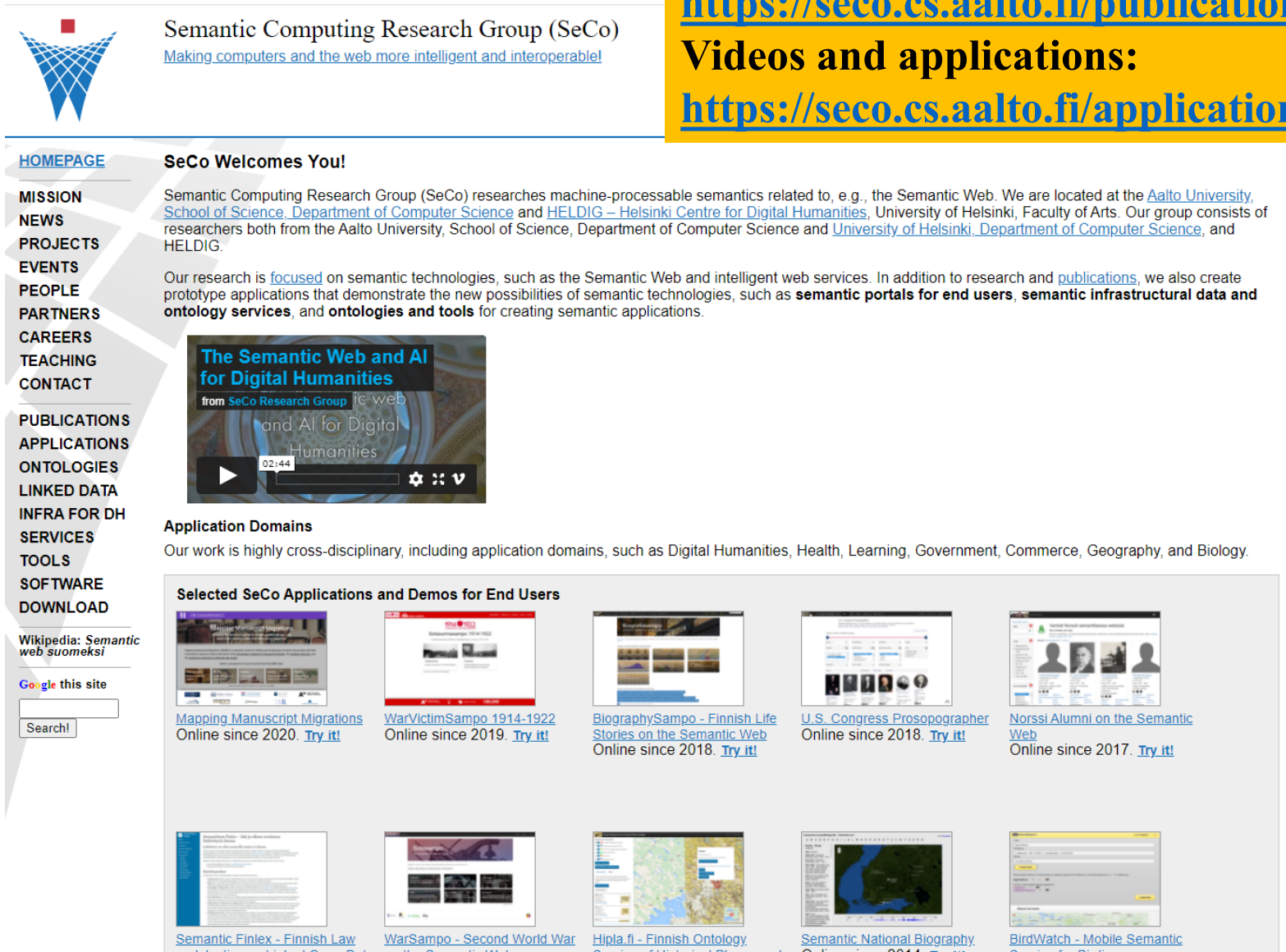
1. Carvings & Printed Texts
2. Online Systems for Searching and Exploring
3. Publishing Content as Linked Data with Tools for DH
4. Automatic Knowledge Discovery and Artificial Intelligence

Eero Hyvönen: **Using the Semantic Web in Digital Humanities: Shift from Data Publishing to Data-analysis and Serendipitous Knowledge Discovery**. Semantic Web, vol. 11, no. 1, pp. 187-193, 2020.

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/



Example Video of a Sampo System: https://vimeo.com/461293952