

Semantic Web in Digital Humanities: Paradigm Shift from Text Publishing to Serendipitous Knowledge Discovery

HELDIG Digital Humanities Summit 2020
University of Helsinki, Helsinki, Dec 11, 2020

Prof. Eero Hyvönen, Director
Helsinki Centre for Digital Humanities (HELDIG)
University of Helsinki and Aalto University
Semantic Computing Research Group (SeCo)
<http://seco.cs.aalto.fi/u/eahyvone/>

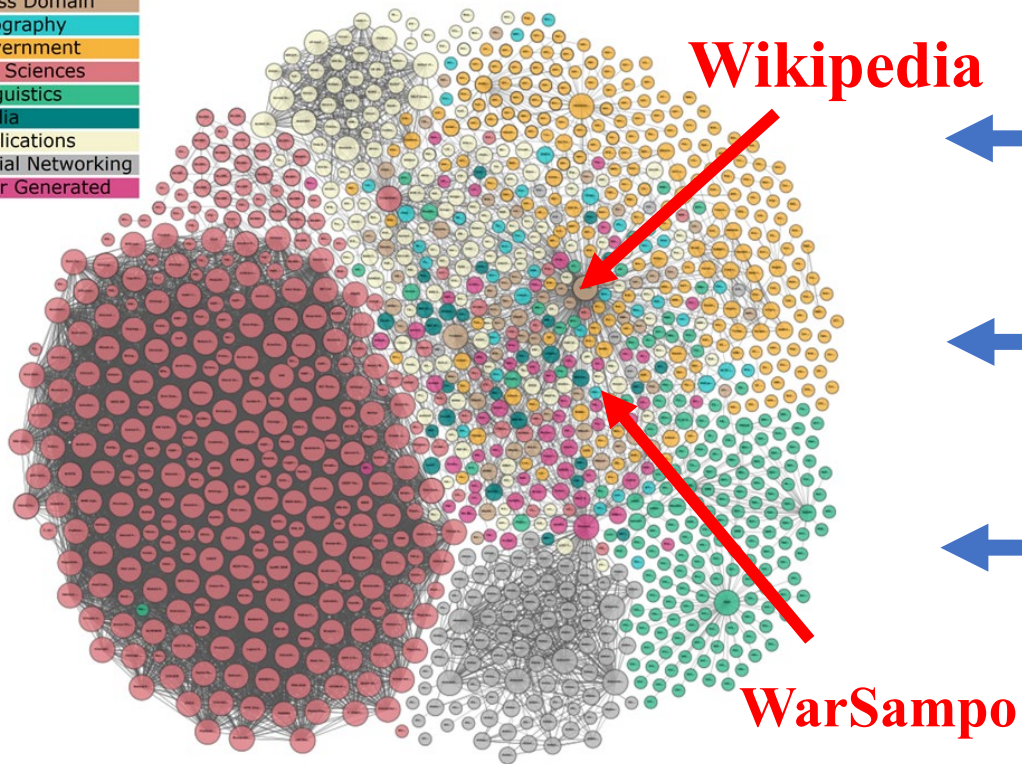
Contents

*Overview of ca 20 years of work in Finland since 2001
at the Semantic Computing Research Group (SeCo) in Aalto Univ. and Univ. of Helsinki*

- **Why Linked Open Data (LOD) for Digital Humanities?**
- **LOD infrastructure for Digital Humanities in Finland**
- **Paradigm Shift from Text Publishing to Knowledge Discovery & Artificial Intelligence**
 - BiographySampo as an example

Linked Data & Semantic Web

Human Knowledge on the Web:
"Web of Data" for machines



Linked Open Data Stats 2018 (<http://stats.lod2.eu/>):
10 000 datasets, 150 billion triples

Traditional Web
"Web of Pages" for humans



Google Index (2018):
50 billion pages (excluding hidden web)

Big Boys Have Entered the Game: Knowledge Graphs

<http://schema.org>

- Google Knowledge Graph
- Microsoft Satori
- IBM Watson
- eBay Products
- Facebook Graph
- ...



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/>



Vision:
**National Linked Data Infrastructure
for Cultural Heritage & Digital Humanities in Finland**



Building a National Level Linked Open Data Infrastructure for Digital Humanities in Finland



<https://vimeo.com/460086143>



Paradigm Shift from Text Publishing to Knowledge Discovery & Artificial Intelligence

BiographySampo as an Example

BiographySampo: Artificial Intelligence Reading Biographies for the Semantic Web



BiographySampo
Artificial Intelligence Reading Biographies for the Semantic Web

BiographySampo - Finnish Biographies on the Semantic Web NOT YET RATED

2 months ago | More

SeCo Research Group

[Settings](#) [Review tools](#) [Send file](#)

This column is only visible to you

See all video stats [View Stats](#)

Related Videos

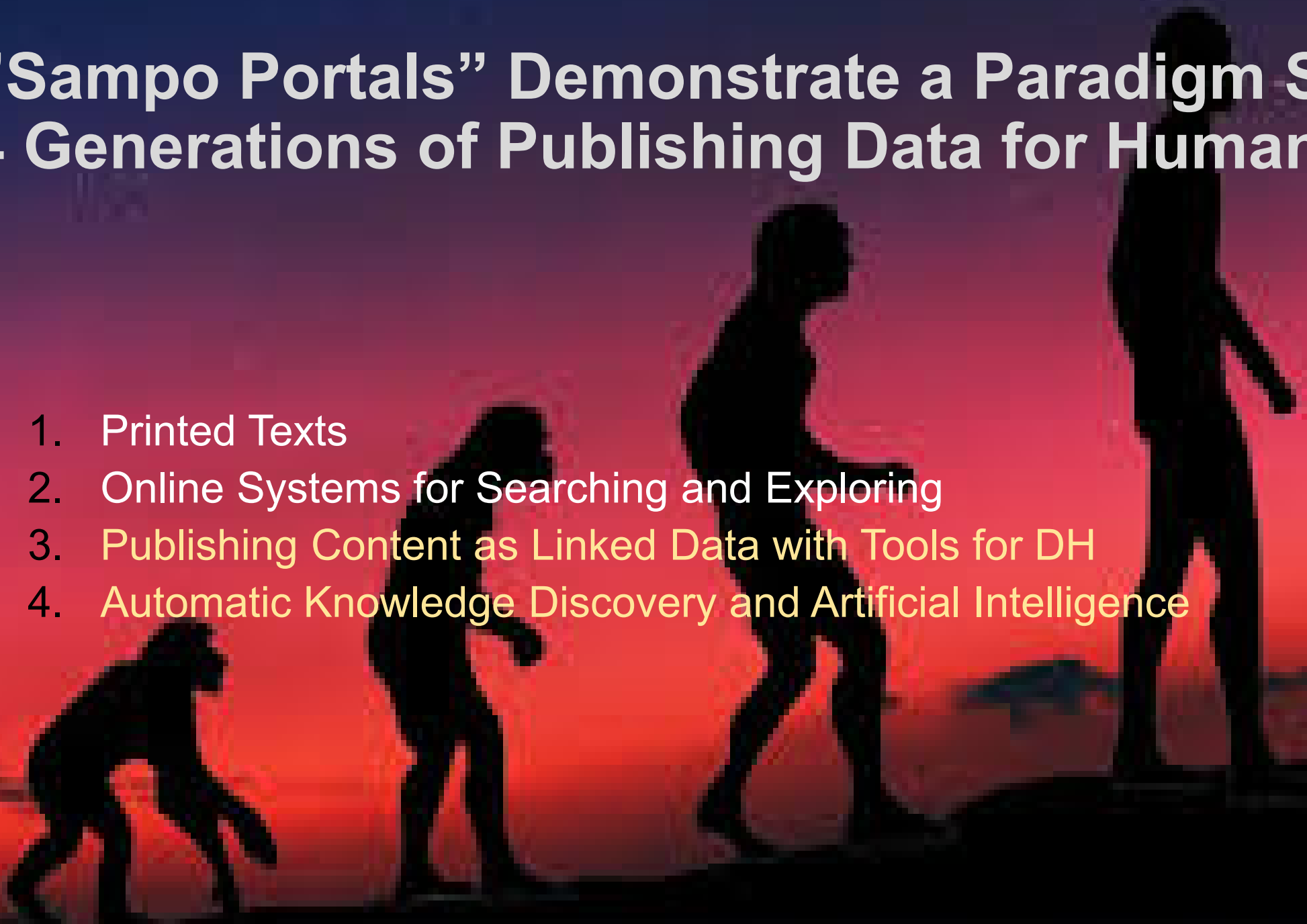
Autoplay next video

BiographySampo - ...
SeCo Research Group

<https://vimeo.com/328419960>

"Sampo Portals" Demonstrate a Paradigm Shift: 4 Generations of Publishing Data for Humanities

1. 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



But the Lunch is not Free

- More collaboration is needed -> complicates work
- Integration of semantic portals with legacy systems
- Manual annotations are costly and may not scale up
- Automatic annotation and linking lowers data quality



Need more source criticism and data literacy!

- What the data actually is and means?
- Big data quality issues: completeness, skewness, errors

More Information – Questions?

Linked Open Infra for Digital Humanities in Finland: LODI4DH

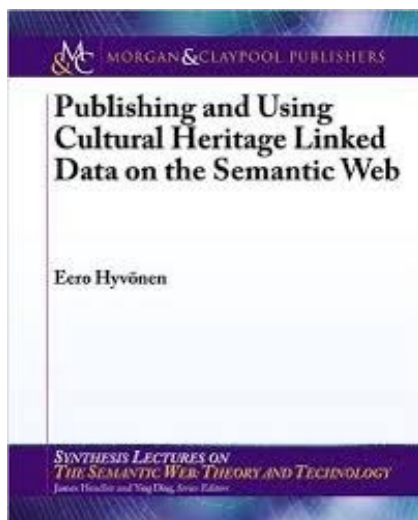
<https://seco.cs.aalto.fi/projects/lodi4dh/>

Sampo Model & Portal Series:

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

Semantic Web & Linked Data Standards

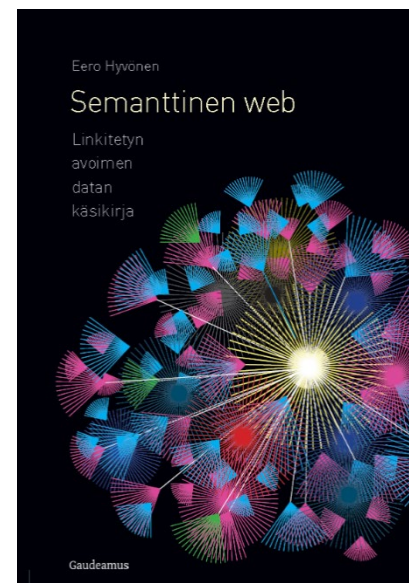
<http://www.w3.org/standards/semanticweb/>



In English

2012

<https://www.amazon.com/Publishing-Cultural-Heritage-Synthesis-Technology/dp/1608459977>



In Finnish

2018

<https://www.gaudemus.fi/semanttinen-web/>