



Sotasampo

Visio, missio ja toteutus

Talvi- ja jatkosota semanttisessa webissä

Prof. Eero Hyvönen

Semantic Computing Research Group (SeCo)

Aalto University, Dept. of Computer Science

<http://www.seco.tkk.fi/>

Outline

- Background
- Vision
- Mission
- Realization

Background

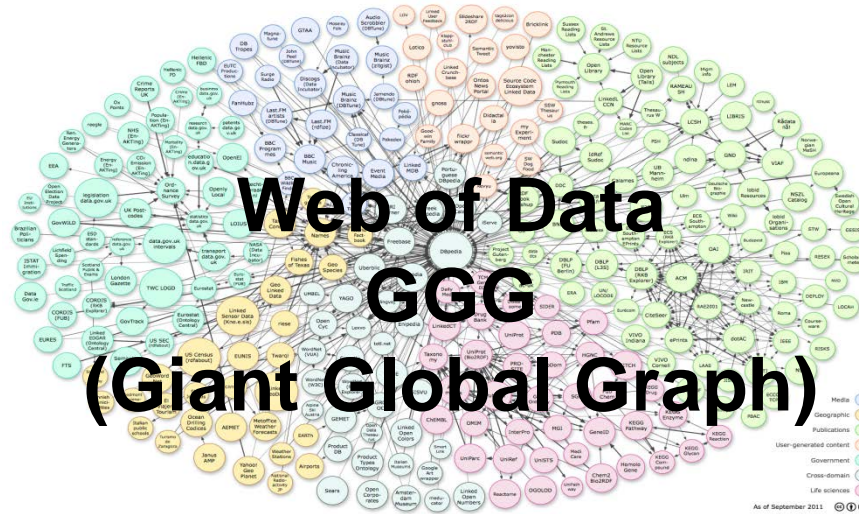


Digital Humanities Research

- Digital Humanities = Computer Science + Humanities
- Major Goals
 - Computer Science Tools for Humanities Research
 - Modelling
 - Analyzing
 - Discovering
 - Applications for the Public
 - E.g. Cultural Heritage portals

Semantic Web and Linked Data Research

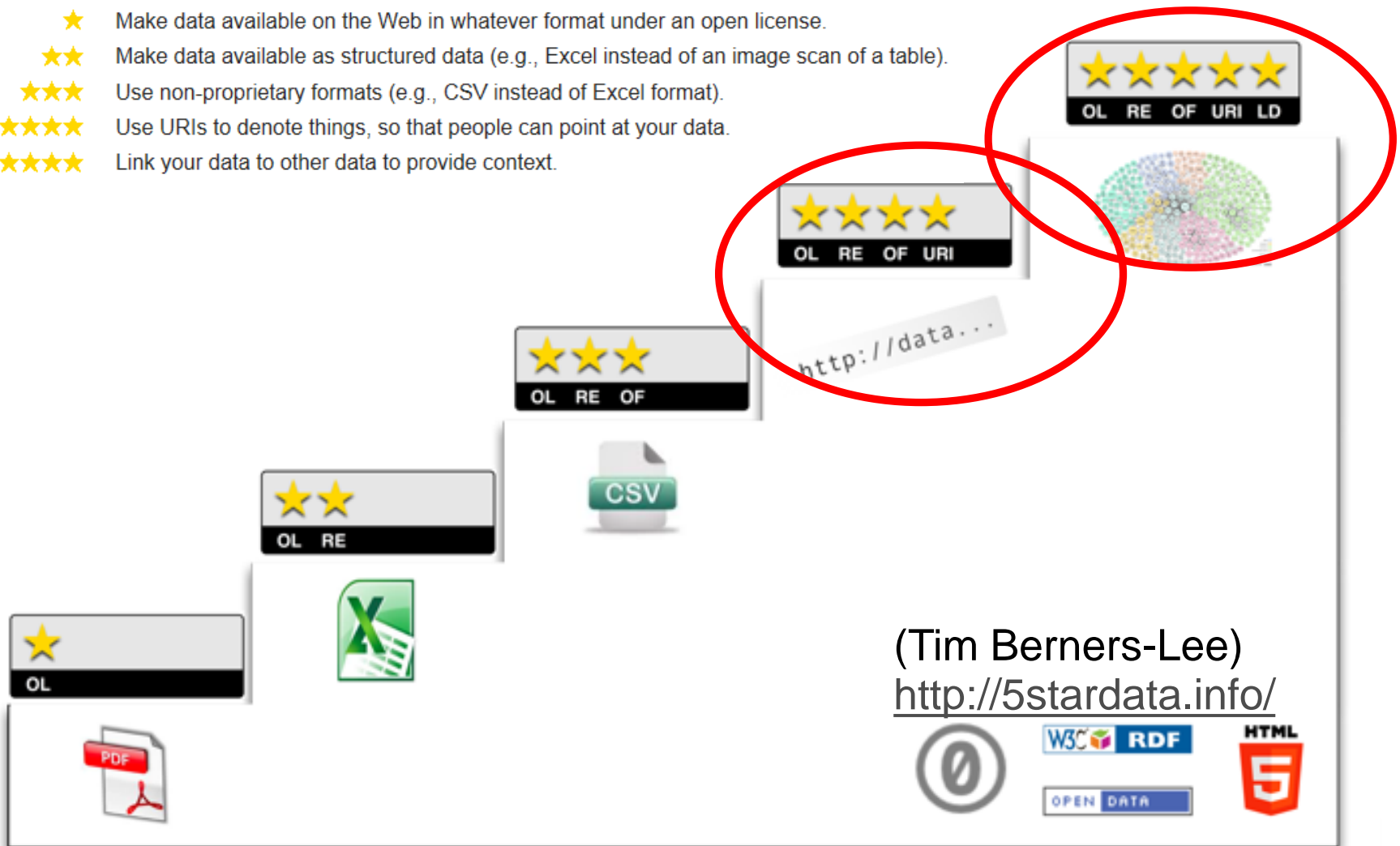
**Web of Pages
WWW**



How to Publish Linked (Open) Data?

5-star Linked Data Model

- ★ Make data available on the Web in whatever format under an open license.
- ★★ 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.





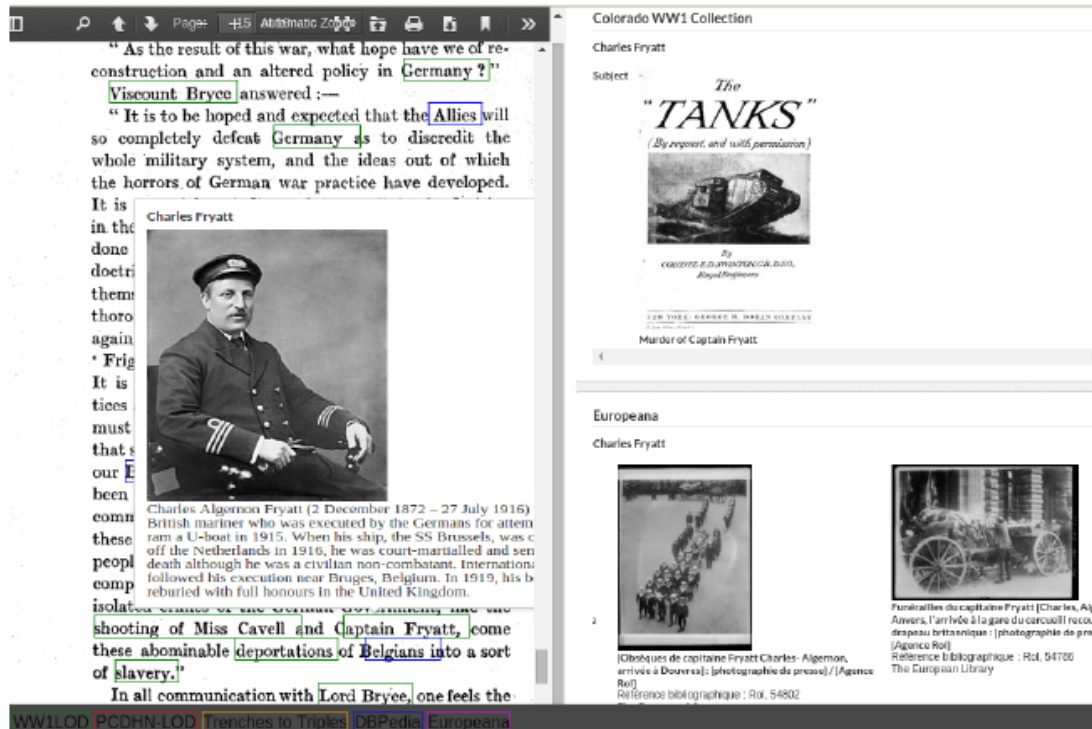
Linked Data Finland

<http://ldf.fi>



World War I on the Semantic Web

- Aalto/SeCo with University of Colorado Boulder
- <http://www.seco.tkk.fi/projects/history>



“As the result of this war, what hope have we of reconstruction and an altered policy in Germany?”
Viscount Bryce answered :—
“It is to be hoped and expected that the Allies will so completely defeat Germany as to discredit the whole military system, and the ideas out of which the horrors of German war practice have developed. It is in the done doctrine them thoro again ‘Frig It is tices must that s our E been comm these peopl comp isolat shooting of Miss Cavell and captain Fryatt, come these abominable deportations of Belgians into a sort of slavery.”
In all communication with Lord Bryce, one feels the

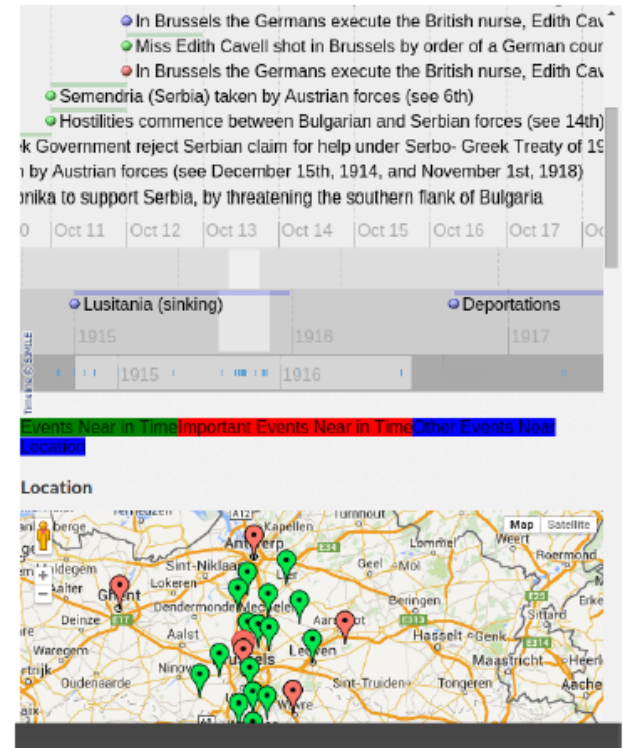
Colorado WW1 Collection
Charles Fryatt
Subject
“The TANKS”
(By request and with permission)
By CHARLES FRYATT, ROYAL ENGINEER
FOR THE LONDON & NORTH OXFORD RAILWAY
Murder of Captain Fryatt

Europeana
Charles Fryatt

Charles Algernon Fryatt (2 December 1872 – 27 July 1916) British mariner who was executed by the Germans for attempting to ram a U-boat in 1915. When his ship, the SS Brussels, was captured off the Netherlands in 1916, he was court-martialed and sentenced to death although he was a civilian non-combatant. International attention followed his execution near Bruges, Belgium. In 1919, his body was reburied with full honours in the United Kingdom.

Funérailles du capitaine Fryatt [Charles, Algernon, arrivé à Douvres] : photographie de presse / [Agence Rol] Référence bibliographique : Rol, 54706 The European Library

(a)



In Brussels the Germans execute the British nurse, Edith Cavell
Miss Edith Cavell shot in Brussels by order of a German court
In Brussels the Germans execute the British nurse, Edith Cavell
Semendria (Serbia) taken by Austrian forces (see 6th)
Hostilities commence between Bulgarian and Serbian forces (see 14th)
Serbian Government reject Serbian claim for help under Serbo-Greek Treaty of 1913
Austrian forces (see December 15th, 1914, and November 1st, 1918)
Austrian forces threaten to support Serbia, by threatening the southern flank of Bulgaria

Oct 11 Oct 12 Oct 13 Oct 14 Oct 15 Oct 16 Oct 17 Oct 18

Lusitania (sinking) Deportations

1915 1916 1917

Events Near in Time Important Events Near in Time Other Events Near in Time

Location

Map of Belgium showing locations of events near in time, important events near in time, and other events near in time.

(b)

Fig. 2. The contextual reader interface

Vision



Goals

- Create a National Open Shared WW2 Data Repository
 - For historians to use the data in research
 - For the public to understand and know Finnish history
 - To promote peace and Finnish identity as a nation
- Demonstrate Applications of the Data
 - To support Digital Humanities research
 - Cultural Heritage publication for citizens
 - To support collaborative Citizen Science
- Develop Semantic Web Technology
 - Data Production and Publication
 - Annotation, enriching, linking, validation, publications
 - Data analysis, visualization, and knowledge discovery

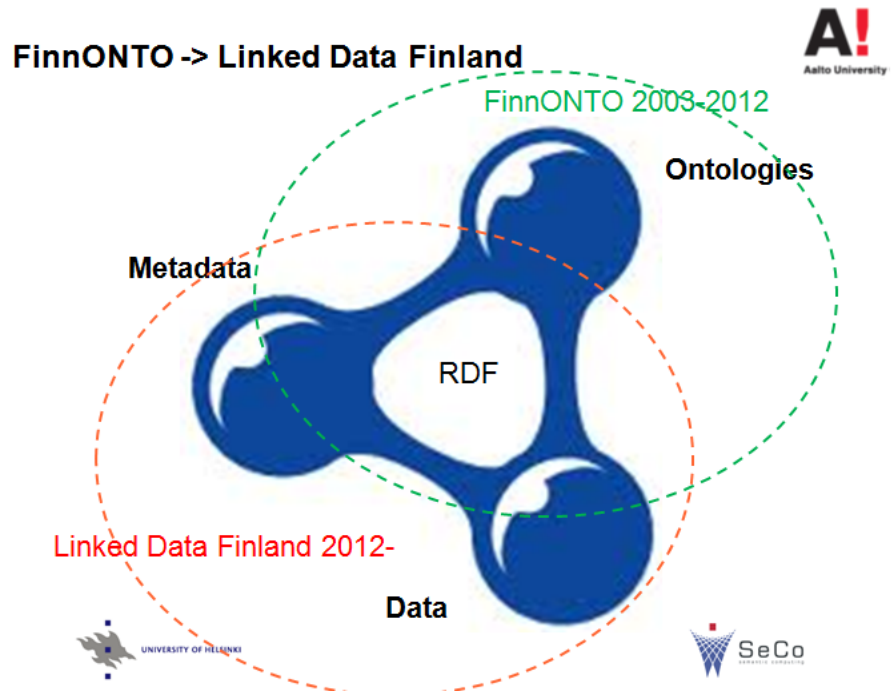


**”We learn from history, that
we learn nothing from history”**

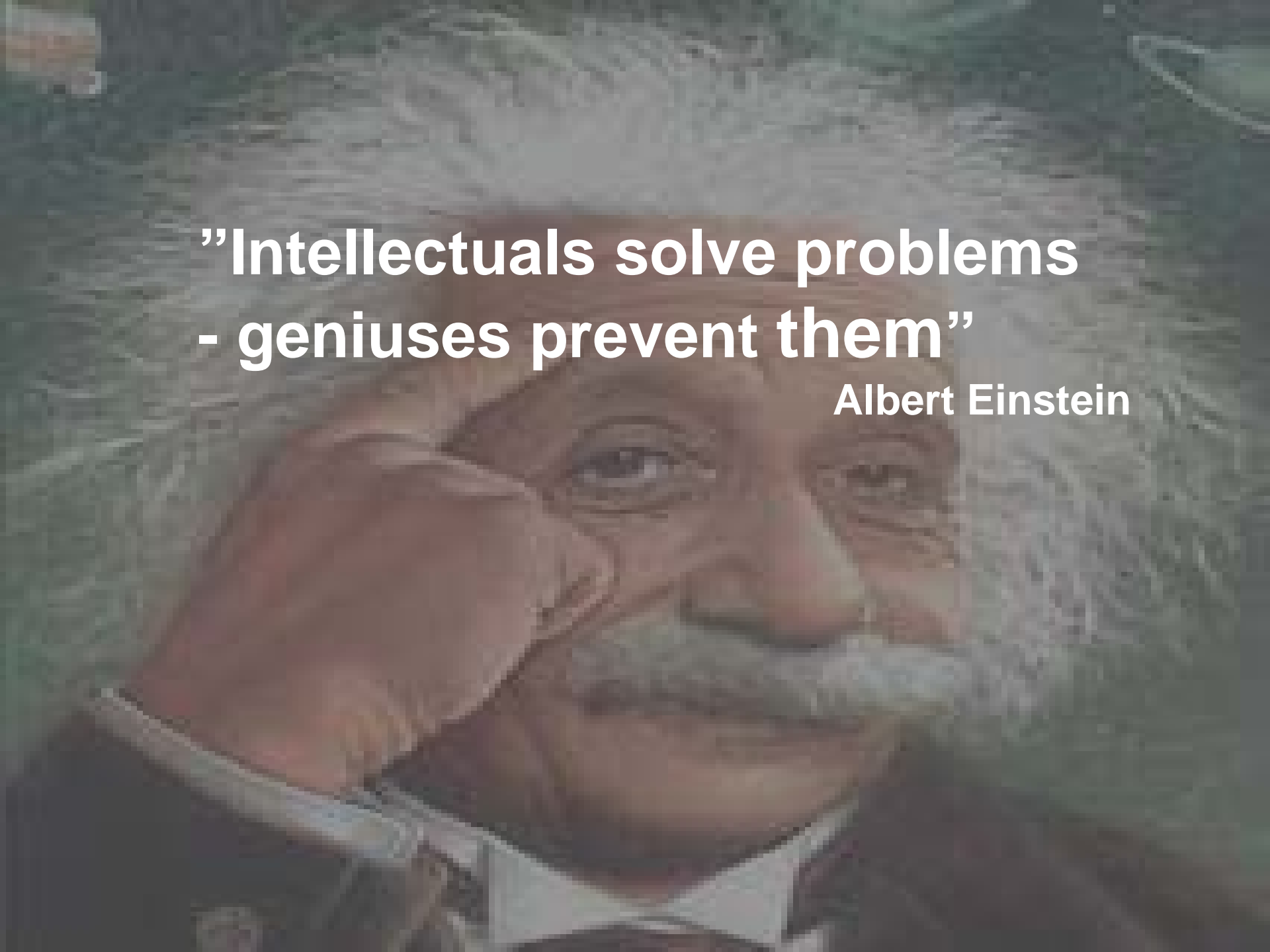
Georg Wilhelm Friedrich Hegel

Preventing Interoperability Problems in Advance by Collaboration

- Sharing **Ontologies**
- Sharing **Metadata Schemas**
- Sharing **Linked Open Data**



[Hyvönen et al., ICSC 2007, ESWC 2009, SW Journal 2010]



**”Intellectuals solve problems
- geniuses prevent them”**

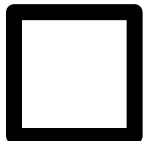
Albert Einstein

Mission



Strategic Steps

1. **Collaboration network.** Create a consortium of organizations and individuals interested in publishing and using Finnish WW2 data
2. **Project and Funding.** Get funding for doing the needed research and development
3. **Prototype.** Test the idea by creating a Living Laboratory prototype for the data publishing system & applications
4. **Sustainable ecosystem.** Establish a collaborative ecosystem for developing and maintaining the data service and its applications



Realization



1. Collaboration Network

- Project Coordination Group

Yritys / organisaatio	Jäsen
Kansallisarkisto	Tomi Ahoranta
Puolustusvoimien Kuvakeskus	Pasi Leskinen
Sotamuseo	Johanna Palokangas
Yleisradio Oy	Pia Virtanen
Suomalaisen Kirjallisuuden Seura SKS	Lauri Harvilahti
Svenska Litteratursällskapet SLS	Dennis Stolt
Opetus- ja kulttuuriministeriö	Sami Niinimäki
Aalto-yliopisto	Eero Hyvönen

- Other Collaborators

- Finnish Association of Military History, National Land Survey, National Defense University, Napafilms Oy, Lapin maakuntamuseo, individual persons

- Future plans

- Extending collaborations to citizen scientists?

2. Project and Funding

- 2013-2014 Concept Development in Aalto
 - Discussions with Collaborating Organizations
- 2014 Work Started with National Archive
 - Published War Casualty Databases as Linked Open Data
- 2015 Part of Linked Open Data Science –project
 - Funded by Ministry of Education and Culture
 - Part of Open Science and Research 2014-2017 Programme
 - A case study in Linked Open Data Science –project of Aalto
- 2015 Part of Centenary of Finland's Independence 2017



WarSampo Research Team at Aalto University, Semantic Computing Research Group (SeCo)



- Eero Hyvönen (leader)
- Jouni Tuominen
- Eetu Mäkelä
- Erkki Heino
- Esko Ikkala
- Mikko Koho
- Petri Leskinen
- Minna Tamper



3. Prototype: two components

Data service for Linked Open Data
<http://ldf.fi>

Applications based on the service
<http://sotasampo.fi>



Linked Data Finland
Living Laboratory Data Service for the Semantic Web

This site is the Living Laboratory of the [Linked Data Finland](#) research initiative, conducted by the [Semantic Computing Research Group](#) at [Aalto University](#) in collaboration with University of Helsinki and a large consortium of Finnish public organizations and companies.

Our goal is to make life easier for both publishers as well as consumers of structured data on the Web. We base our work on the [Linked Data](#) paradigm and stack of standards, which combines an expressive, semantic data model ([RDF](#)) with standardized access mechanisms ([SPARQL](#) and [live HTTP URIs](#)).

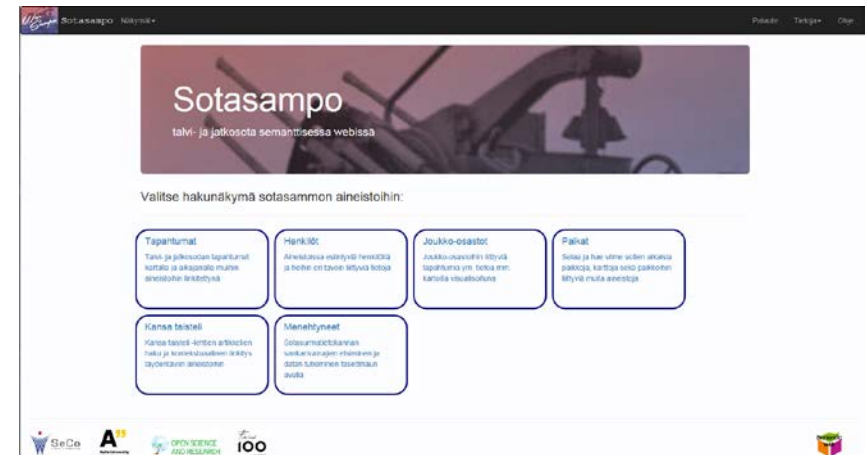
5-star Linked Data

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

7-star Linked Data Service

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 and access them, as well as a better grasp of their quality and provenance. To this end, we extend the model with two additional stars.



Sotasampo
talti- ja jatkosota semanttisessä webissä

Välitse hakukymä sotasammon aineistoihin:

- Tapahtumat**
Tavik- ja jalkaväen tapahtumat, koulutus ja aikapöytä muutkin aineistot katsottavissa
- Henkilöt**
Aineistossa esitetyt henkilöt ja heidän on tavallaan löydettävissä
- Joukko-osastot**
Aika- ja paikkatietoisia löytyviä tapahtumia ym. tietoja mm. kartoista vuosituloksesta
- Paiikat**
Sotaa ja sen välineiden aikana paikoissa, karttoja sekä paikkoista löytyviä muita aineistoja
- Käsiä taiteella**
Käsiä taiteella -lehden esittelyt sekä ja verkkokuvalehden arkistot ja yhtenäiset aineistot
- Mennehtyneet**
Sotasuomalaisenaan kuolleiden ja vammautuneiden ja sotien tuomien sotavainajien

Logos for SeCo, Aalto University, OPEN SCIENCE, and i100 are visible at the bottom.



WarSampo

Linked Data Finland



[Home](#)

[Project](#)

[Datasets](#)

[Schemas](#)

[Services](#)

[Policies](#)

[Documentation](#)

[Validation](#)

[Applications](#)

[Your Data?](#)

This dataset includes harmonized data of different kinds concerning the Second World War in Finland, separated in different graphs representing events, actors, places, photographs, and other aspects and documentation of the war. To test and demonstrate its usefulness, this data service is in use in the semantic portal [WarSampo](#) explained in more detail in this [project page](#).

License

[CC BY 4.0](#)



See possible graph-specific licenses below.

Detailed Dataset Contents

Wartime Events (URI: <http://ldf.fi/warsa/events>)



([Browse data](#) / [View in Sotasampo.fi](#) / [Download](#))

Information on the events of the Winter War and Continuation War were derived from the books Talvisodan pikkujättiläinen, Jatkosodan pikkujättiläinen, and the organization records of the Finnish Defence Force.

Information included: description, time span, place, participants, type.

The primary schema used to model the data is the [CIDOC-CRM](#).

Semantic Portal

<http://sotasampo.fi>

- Six perspectives to the same data
- Only SPARQL API is used to access data

The screenshot shows the Sotasampo Semantic Portal interface. At the top, there is a navigation bar with the Sotasampo logo and the text "Näkyvät". On the right side of the navigation bar, there are links for "Palautte", "Tietoja", and "Ohje". Below the navigation bar is a large banner image with the text "Sotasampo" and "taivi- ja jatkosota semanttisessa webissä". Below the banner, there is a section titled "Valitse hakunäkymä sotasammon aineistoihin:" (Select search view for Sotasampo's datasets:). This section contains six buttons, each representing a different search perspective:

- Tapahtumat**: Taivi- ja jatkosodan tapahtumat kartalla ja aikajanaa muihin aineistoihin linkitettyinä
- Henkilöt**: Aineistoissa esiintyviä henkilöitä ja heihin eri tavoin liittyviä tietoja
- Joukko-osastot**: Joukko-osastoihin liittyviä tapahtumia ym. tietoja mm. kartoilta vuositsoituna
- Paikat**: Selaa ja hae viime soden aikaisia paikkoja, karttoja sekä paikkoihin liittyviä muita aineistoja
- Kansa taisteili**: Kansa taisteili -lehden artikkeleiden haku ja kontekstuaalinen linkitys täydentäviin aineistoihin
- Menehtyneet**: Sotasurmatutokannan sankarivainajien etsiminen ja datan tutkinnin laajentaminen avulla

At the bottom of the page, there are several logos: SeCe, A! Aalto University, OPEN SCIENCE AND RESEARCH, Finland 100, and a small logo on the right.

Linked Open Datasets in Sotasampo

(Nov 22, 2015)

Aineisto (25.11.2015)	Määrä	Datan lähteet ja tekijänoikeus	Lisenssi linkitettyyn dataan
TK-kuvat ja videot	160 000	Puolustusvoimat & Aalto/SeCo	CC BY 4.0
Menehtyneet 39-45	95 000	Kansallisarkisto & Aalto/SeCo	CC BY 4.0
Organisaatiokortisto	500	Kansallisarkisto & Aalto/SeCo	CC BY 4.0
Sotapäiväkirjat	26 500	Kansallisarkisto & Aalto/SeCo	CC BY 4.0
Senaatin kartat	430	Kansallisarkisto & Aalto/SeCo	CC BY 4.0
Kansa taisteli -lehtien artikkelit 1957-1986	3 360	Metadata: Timo Hakala & Aalto/SeCo. Artikkelit © Suomen Sotahistoriallinen Seura ja Bonnier, http://sshs.fi .	CC BY 4.0 (metadata)
Paikannimirekisteri	800 000	Maanmittauslaitos (MML) & Aalto/SeCo	CC BY 4.0
Karjalan topograf. kartat	46	MML & Aalto/SeCo	Ei-kaupallisiin tarkoituksiin
Karjalan paikat geokood.	35 000	Jyrki Tiittanen & Aalto/SeCo	CC BY 4.0
Sotien tapahtumat	11 275	Aalto/SeCo. Lähteet: Talvi- ja jatkosodan pikkujättiläiset, organisaatiokortisto ym.	CC BY 4.0
Henkilöt (Menehtyneet 39-45 datan lisäksi)	1 050	Aalto/SeCo. Lähteet: Talvisodan historia, sotapäiväkirjat, organisaatiokortisto, Suomen Rintamamiehet 1939-1945, Maanpuolustuskorkeakoulu/Ohto Manninen, Wikipedia ym.	CC BY 4.0
Biografiat	6 300	Suomalaisen Kirjallisuuden Seura / Kansallisbiografia & Aalto/Seco	CC BY 4.0
Joukko-osastot	3 180	Aalto/SeCo. Talvisodan historia, Talvisodan sotapäiväkirjat, organisaatiokortisto	CC BY 4.0
	1 142 641		

Technological Basis

- Key goals/ideas of the design
 - **Enriching** each others' data by data linking
 - **Reusing** the shared data in applications
 - **Intelligent applications** based on semantic computing
- Data Service
 - Linked Data publishing principles
 - **Ontologies** as "semantic glue" (actors, places, events, times,...)
 - Event-based **metadata** harmonization (CIDOC CRM)
 - **Linked Open Data** (RDF, OWL)
- Sotasampo.fi applications
 - Rich Internet Application (RIA) principles (HTML5, Javascript, ...)
 - Only dataservice on the server side

Summary



Open your data so that it links with the others' data!

- 1. Everybody's **data is enriched** for "free"
- 2. Redundant **work is minimized** by collaboration
- 3. **Work can be shared** in better ways
- 4. **Data can be reused** in different applications

Use Shared Infrastructure to Prevent Data Linking Problems!



Thanks!

More info and publications on the web:

<http://www.seco.tkk.fi/projects/sotasampo>




Semantic Computing Research Group (SeCo)
Making computers and the web more intelligent and interoperable!

HOME PAGE
MISSION
NEWS
PROJECTS
EVENTS
PEOPLE
PARTNERS
CAREERS
TEACHING
CONTACT
PUBLICATIONS
APPLICATIONS
ONTOLOGIES
LINKED DATA
SERVICES
TOOLS
SOFTWARE
DOWNLOAD

Sotasampo:
Itsenäisen Suomen sotahistoria linkitettyä datana
Tavoitteena historian ymmärtäminen

G. W. F. Hegelin mukaan historia opettaa meille, että emme opi historiasta mitään. Toivottavasti näin ei tapahdu Suomessa toisen maailmansodan osalta, johon liittyvät asiat ovat taas olleet esillä mm. Ukrainan kriisin yhteydessä. Itsenäisyytemme ajan sotahistoria, erityisesti talvi- ja jatkosota, ovat 100-vuotisjuhlinsa valmistautuvan Suomen kansallisen identiteetin ja maanpuolustustahdon keskeinen tukipilari ja laajaa kiinnostusta herättävä historian tutkimuskohde. Sotasampo-hankkeen tavoitteena on rakentaa suomalaista identiteettiä lisäämällä ymmärrystämme viime sodistamme ja edistää näin rauhaa. Työ on jatkoa SeCo-ryhmän aiemmalle tutkimukselle [historiallisen tiedon julkaisemiseksi semanttisessa webissä](#).



Wikipedia: *Semantic* Kuva: Mäkiluodon tykeillä ammutaan Hangon taisteluissa 1942, Puolustusvoimien SA-kuva-arkisto

Latest News
[Linked Open Data Science project starts](#)
A new project starts at SeCo, where a data service for publishing...
4.5.2015 10:11 by eahyvone
[Linked Data Finland Double Winner in Apps4Finland Competition](#)
Apps4Finland – Doing Good with Open Data is a major Finnish...
4.12.2014 13:57 by eahyvone
[State Award for Public Information](#)
Professor Eero Hyvönen receives State Award for Public Information...
25.9.2014 15:40 by eahyvone

Latest Publications
Eetu Mäkelä, Juha Törnroos, Thea Lindquist and Eero Hyvönen: [World War 1 as Linked Open Data](#)
Eero Hyvönen, Jouni Tuominen, Eetu Mäkelä, Jérémie Dutruit, Kasper Apajalahti, Erkki Heino, Petri Leskinen and Esko Ikkala: [Second World War on the Semantic Web: The WarSampo Project and Semantic Portal](#)
Eero Hvvönen, Jouni Tuominen, Esko Ikkala and

Applications: <http://sotasampo.fi/>
Data service: <http://ldf.fi/>

Ohjelma

Aika	Esityksen aihe	Esiintyjä
09:00-09:10	Tervetuloa Kansallisarkistoon	Jaana Kilkki Kansallisarkisto
09:10-09:30	Sotasampo: visio, missio ja toteutus	Eero Hyvönen Aalto-yliopisto
09:30-09:50	Sotasammon datapalvelu ja sen käyttö	Jouni Tuominen Aalto-yliopisto
09:50-10:10	Sovellus 1: Talvi- ja jatkosodan tapahtumat	Erkki Heino Aalto-yliopisto
10:10-10:30	Sovellus 2 ja 3: Henkilöt ja joukko-osastot	Petri Leskinen Aalto-yliopisto
10:30-11:00	Tauko	
11:00-11:20	Sovellus 4: Historialliset paikat ja kartat	Esko Ikkala Aalto-yliopisto
11:20-11:40	Sovellus 5: Sotasurmat	Mikko Koho Aalto-yliopisto
11:40-12:00	Sovellus 6: Kansa taisteli -lehdet ja Semanttinen selain	Minna Tamper ja Eetu Mäkelä Aalto-yliopisto
12:00	Tilaisuuden päätös	