



Aalto-yliopisto
Teknillinen korkeakoulu

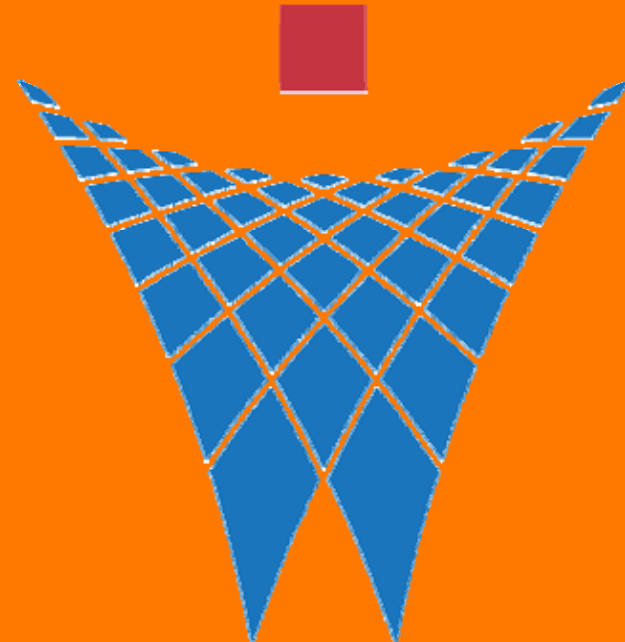
Miten Linked Data – aineistoja tuotetaan ja julkaistaan?

Kim Viljanen

Semanttisen laskennan tutkimusryhmä SeCo

Aalto-yliopisto

kim.viljanen@tkk.fi



Webin ”killer app” on web itse
= valtava määrä informaatiota, joka löytyy
yhdestä osoitteesta ja jossa sivujen
välillä on linkkejä

The Google logo is centered on the page. It consists of the word "Google" in its signature multi-colored font: 'G' is blue, 'o' is red, 'o' is yellow, 'g' is green, 'l' is blue, and 'e' is red. A small trademark symbol (TM) is located to the upper right of the 'e'.

[Advanced Search](#)
[Language Tools](#)

Google Search

I'm Feeling Lucky

**Voiko saman ilmiön toistaa
tietokannoille ja -kokoelmille,
joita maailma on pullollaan?**

**Siten, että tieto olisi
yhteismitallisessa,
yhteenlinkittyneessä muodossa
sovellusten käytettävissä?**

Esimerkiksi: miten Suomen museoiden datat tulisi julkaista muiden käyttöön?



Espoo



Vaasa



Turku

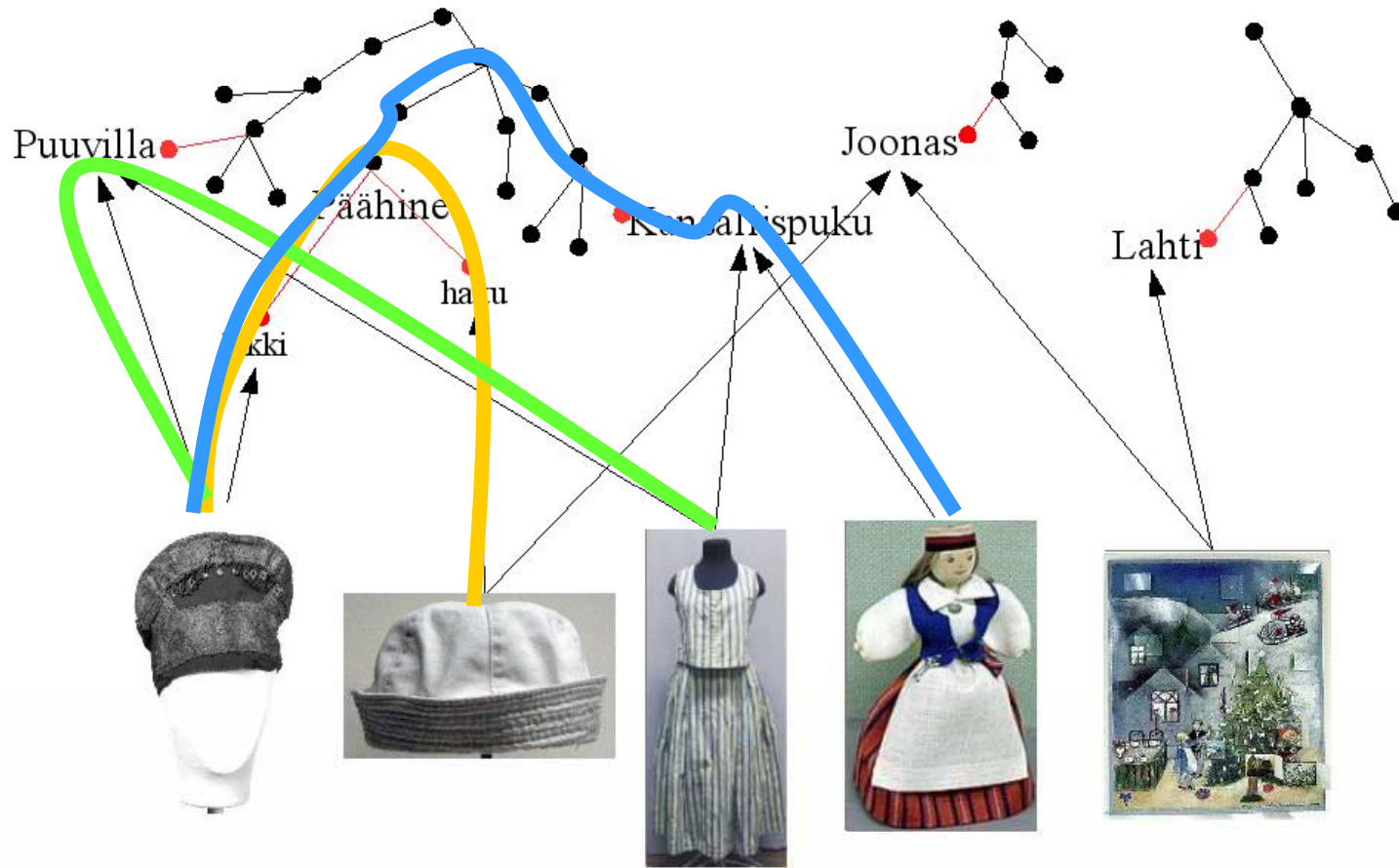


Oulu



Kuopio

...siten että ne rikastavat toisiaan?



Espoo

Vaasa

Turku

Oulu

Kuopio

...siten että syntyisi uusia sovelluksia tämän aineiston päälle?

Kulttuuri-
sampo

X

Y



Espoo



Vaasa



Turku



Oulu



Kuopio

PM Gordon Brown, 22.3.2010

”the next generation of the web is called the semantic web, or the **web of linked data**.

[it] is **a simple concept**, but has the potential to be just as revolutionary – just as disruptive to existing business and organisational models – as the web was itself.”


”Speech on Building Britain’s Digital Future”
<http://www.number10.gov.uk/Page22897>

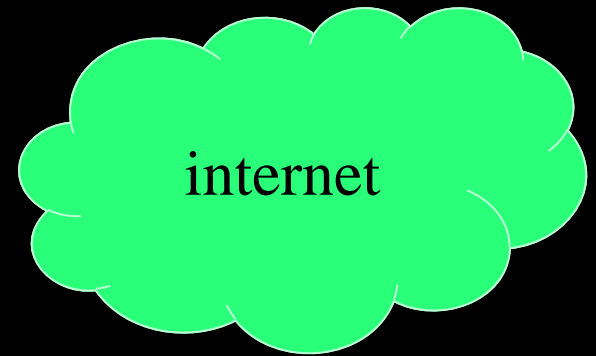
Miten datat julki? Kaksi tasoa.

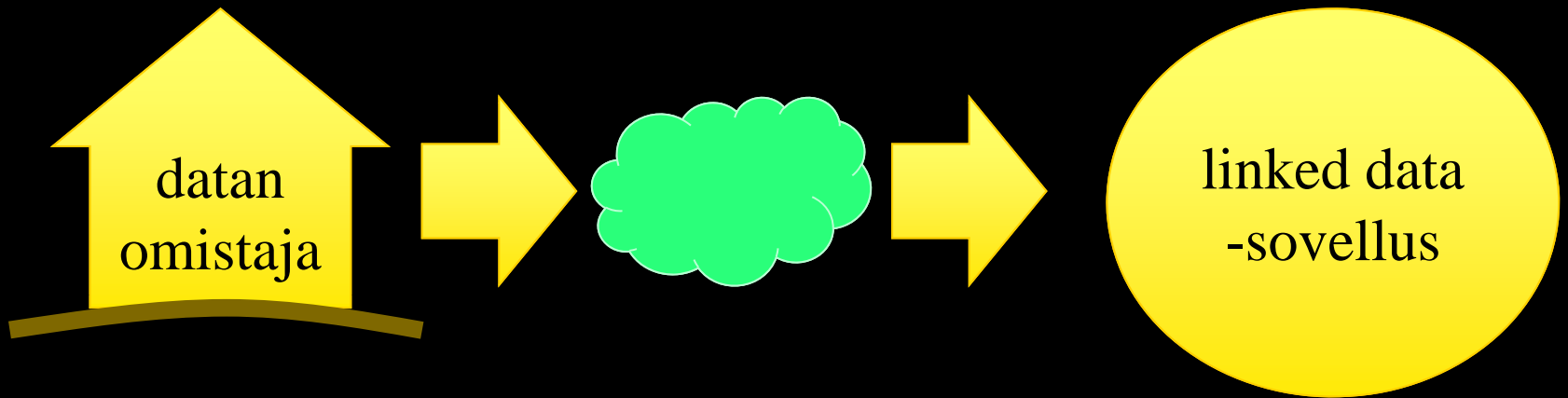


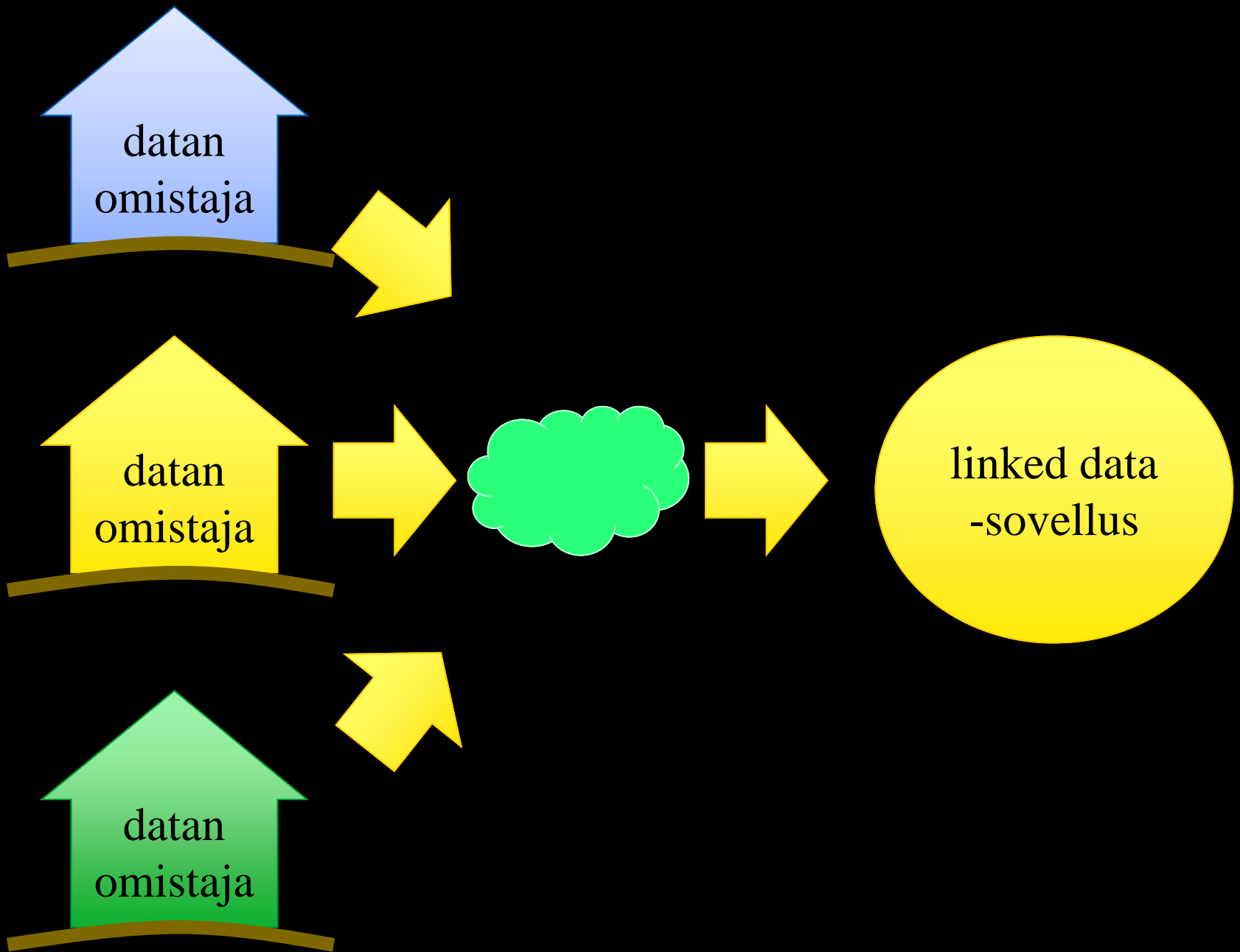
Ihmisten taso

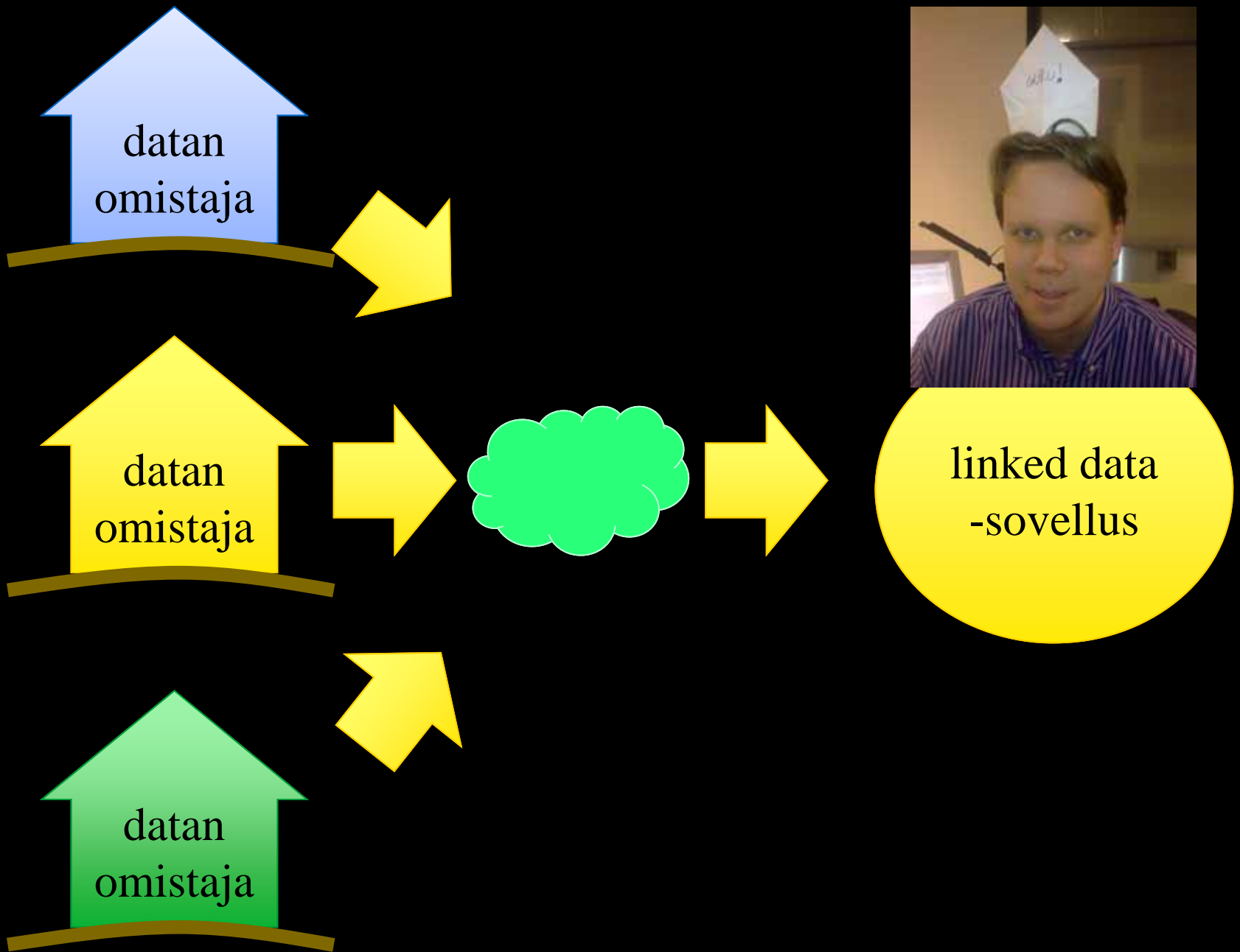
GURU!

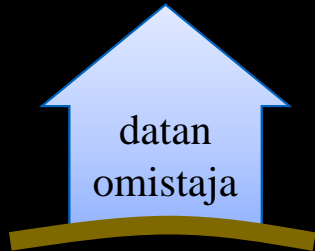
A close-up photograph of a person's head from the top down. They are wearing a white, five-pointed paper hat. The word "GURU!" is written in black marker on the top surface of the hat. The person has dark, wavy hair. The background is a plain, light-colored wall.







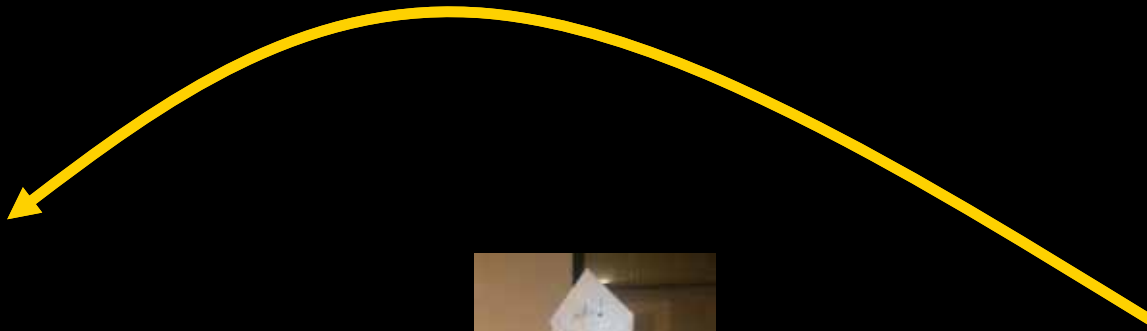


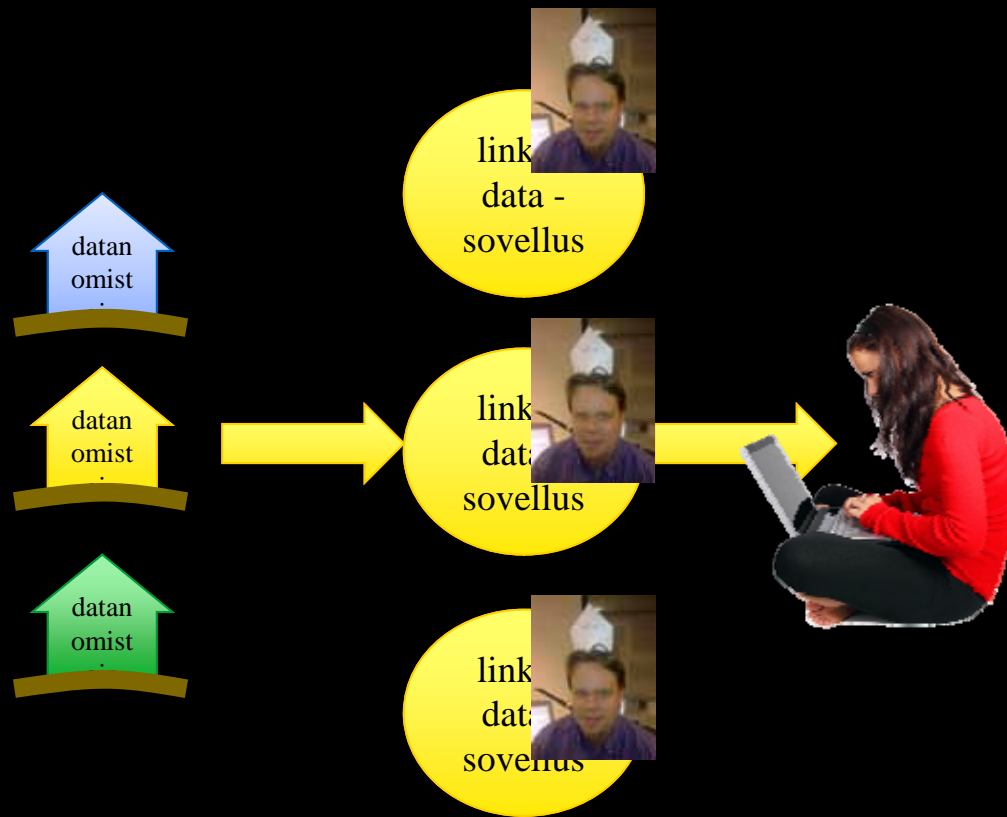


linked
data -
sovellus



loppukäyttäjä





Jotta sovelluksia alkaisi syntymään:

- kehittäjäyhteisön aktivointi
- data houkuttelevaan muotoon
- rahaa, resursseja, ...
- kysyntää

Tekninen taso

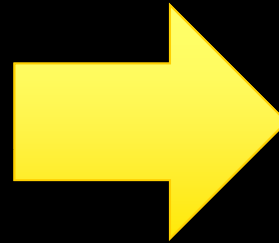


Missä muodossa data julki:

- PDF?
- Excel?
- HTML?
- XML?
- JSON?
- Mikroformatit?
- SQL-rajapinta?
- Web Service?
- ...?

Missä muodossa data julki:

- PDF?
- Excel?
- HTML?
- XML?
- JSON?
- Mikroformatit?
- SQL-rajapinta?
- Web Service?
- ...?



Linked Data

Seitsemän syytä semanttiseen webiin^[1]

1. Parempaa palvelua ihmisille

- semanttinen tiedonhaku, semanttinen suosittelu, jne.

2. Yhteentoimivat järjestelmät

- järjestelmät yhteismitallisiksi merkitystasolla

3. Monikäyttöinen sisältö

- sama semanttinen sisältö monessa sovelluksessa
- yleistää XML:n monikanavajulkaisun sisällölliselle tasolle

4. Joustava sisältöjen yhdistäminen

- eri sisällöt samassa sovelluksessa

5. Hajautettu sisällöntuotanto

- tuotanto voidaan hajauttaa ja sisältö yhdistää semantiikan avulla

6. Kieli- ja kulttuurirajojen ylitys

- käsitteet pitkälti samoja kielestä riippumatta, esim. *tuoli*

7. Web-palveluiden julkaisu & hajautettu hyödyntäminen

- esim. valtionhallinnon portaalit voivat käyttää toisiaan

[1] Eero Hyvönen, 2006





Tim Berners-Lee

Date: 2006-07-27, last change: \$Date: 2009/06/18 18:24:33 \$

Status: personal view only. Editing status: imperfect but published.

[Up to Design Issues](#)

Linked Data

The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data.

Like the web of hypertext, the web of data is constructed with documents on the web. However, unlike the web of hypertext, where links are relationships anchors in hypertext documents written in HTML, for data they links between arbitrary things described by RDF. The URIs identify any kind of object or concept. But for HTML or RDF, the same expectations apply to make the web grow:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs, so that they can discover more things.

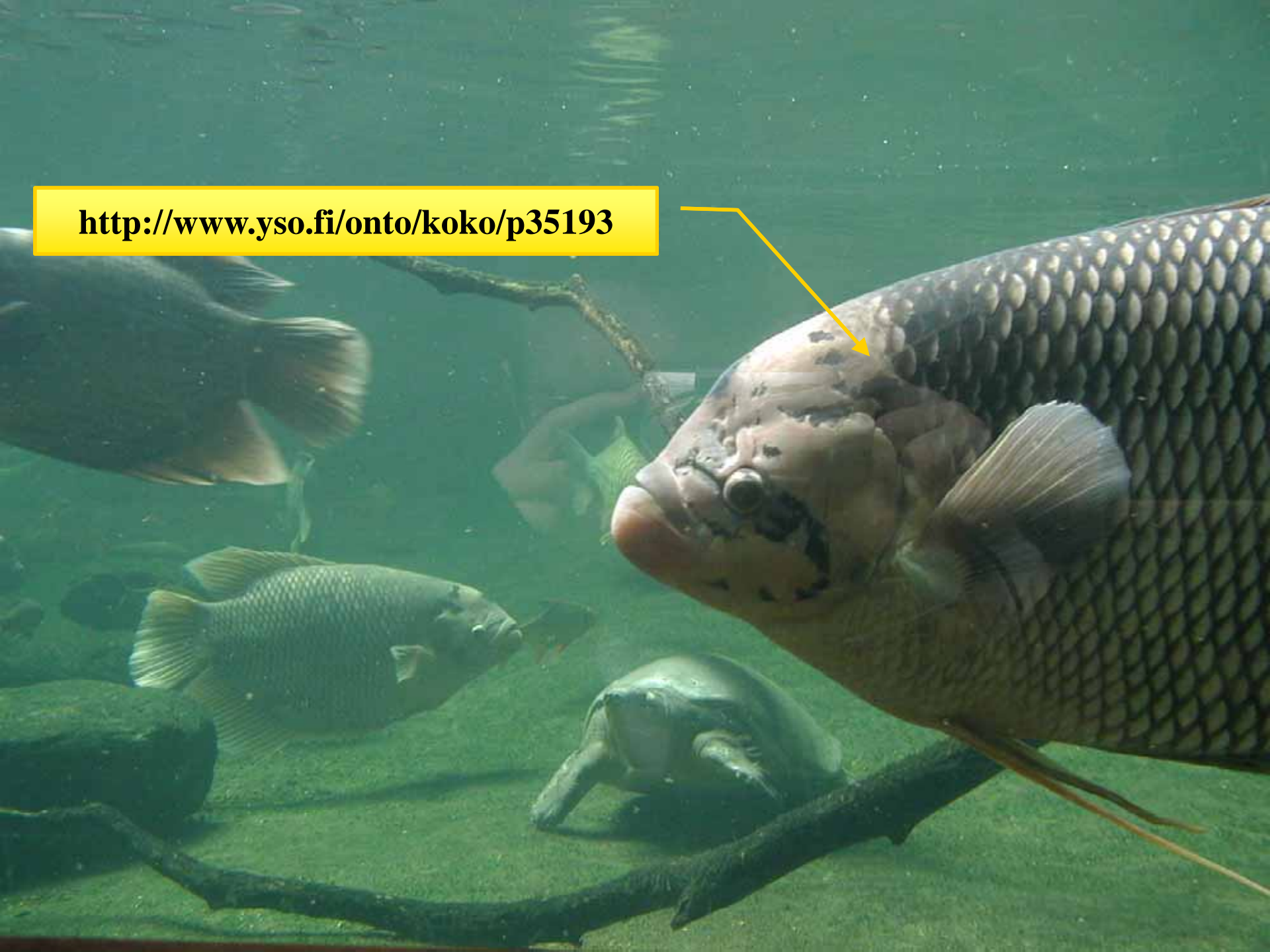
Simple. In fact, though, a surprising amount of data isn't linked in 2006, because of problems with one or more of the steps. This article discusses solutions to these problems, details of implementation, and factors affecting choices about how you publish your data.

The four rules

I'll refer to the steps above as rules, but they are expectations of behavior. Breaking them does not destroy anything, but misses an opportunity to make data interconnected. This in turn limits the ways it can later be reused in unexpected ways. It is the unexpected re-use of information which is the value added by the web.

The first rule, to identify things with URIs, is pretty much understood by most people doing semantic web technology. If it doesn't use the universal URI set of symbols, we don't call it Semantic Web.

<http://www.yso.fi/onto/koko/p35193>





<http://www.yso.fi/onto/koko/p58>

<http://www.kalatehdas.fi/purkki12345>



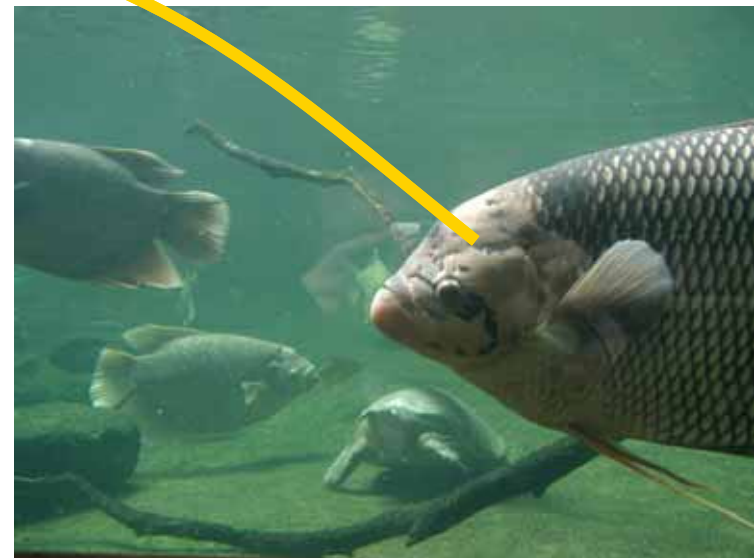
sisältää

<http://www.kalatehdas.fi/purkki12345>

<http://www.yso.fi/onto/koko/p58>

eräänlainen

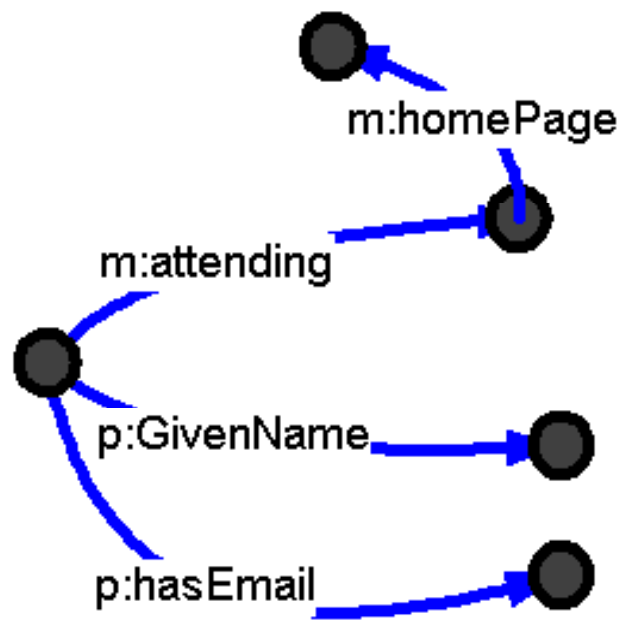
<http://www.yso.fi/onto/koko/p35193>

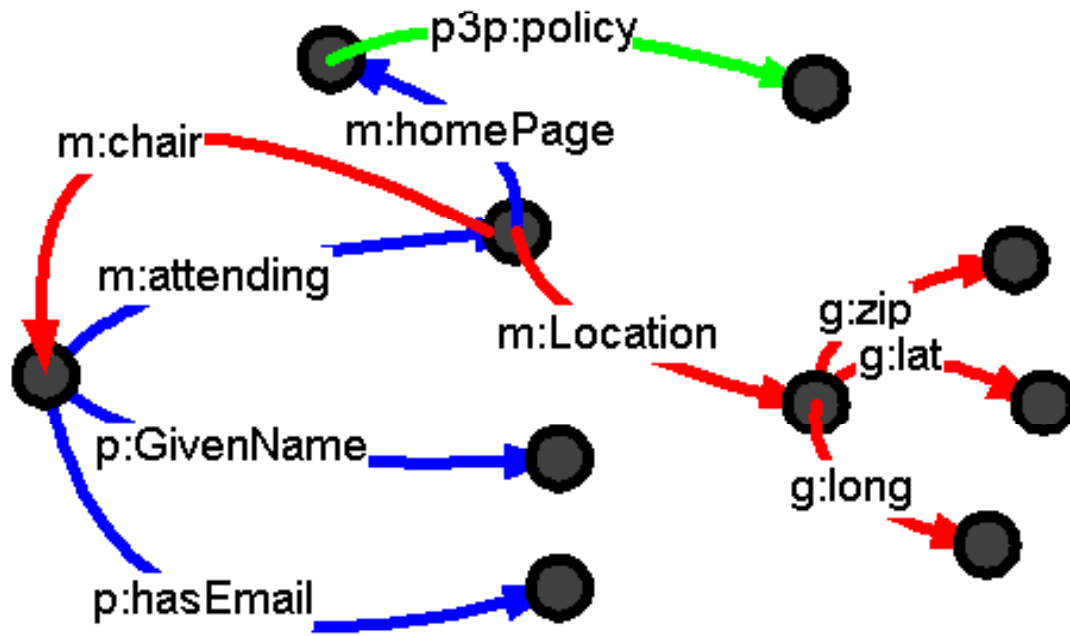




Kätevä kieli: RDF

(Resource description framework)







How to Publish Linked Data on the Web

Authors:

[Chris Bizer](#) (Web-based Systems Group, Freie Universität Berlin, Germany)

[Richard Cyganiak](#) (Web-based Systems Group, Freie Universität Berlin, Germany)

[Tom Heath](#) (Knowledge Media Institute, The Open University, Milton Keynes, UK)

This version:

<http://sites.wiwiss.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/20070727/>

Latest version:

<http://sites.wiwiss.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/>

Abstract

This document provides a tutorial on how to publish Linked Data on the Web. After a general overview of the concept of Linked Data, we describe several practical recipes for publishing information as Linked Data on the Web.

Table of Contents

1. [Introduction: Linked Data on the Web](#)
2. [Basic Principles](#)

1. Linked Data –periaatteet

- kohteiden yksilöinti: URI
- esityskieli: RDF (tai RDFa)



page

discussion

view source

history

RDF



Resource

- Publication date: 2004-02-10 (with a previous version published at: 1999-02-22)
- Created by: [RDF Working Group](#)
- List of documents at: <http://www.w3.org/standards/techs/rdf>

Description

Framework (RDF)

Overview

RDF is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed.

RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a "triple"). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications.

This linking structure forms a directed, labeled graph, where the edges represent the named link between two resources, represented by the graph nodes. This *graph view* is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations.

Recommended Reading

The RDF Specification consists of a [suite of W3C Recommendations](#), published in 2004. This suite also includes an [RDF Primer](#). See also Tim Berners-Lee's writings on [Web Design Issues](#), including [Metadata Architecture](#). Other technologies, like [OWL](#) or [SKOS](#), build on RDF and provide language for defining structured, Web-based ontologies which enable richer integration and interoperability of data among descriptive communities.

A number of textbooks have been published on RDF and on Semantic Web in general. Please, refer to a [separate page](#) listing some of those, as maintained by the community. That list also includes references to conference proceedings and article collections that might be of general interest.

navigation

- [Main Page](#)
- [Recent changes](#)
- [Tools](#)
- [Books](#)
- [Validators](#)

other w3c resources

- [Activity news](#)
- [Publications](#)
- [Logos, buttons](#)
- [Activity home page](#)

w3c rss feeds

- [Activity news](#)
- [W3C blogs](#)
- [Use cases, case studies](#)

account request

- [W3C Member](#)
- [Public](#)

search

Go

Search

toolbox

- [What links here](#)
- [Related changes](#)
- [Upload file](#)



Semantic Computing Research Group (SeCo)

Making computers and the web more intelligent and interoperable!

- » [print this page!](#)
- » [Follow us on Twitter](#)
- » [Fan us on Facebook](#)
- » [0 Comments](#)

>>> [Yhdistetty avoin tieto Suomessa. Perjantaina 26.3.2010. Ilmoittaudu nyt!](#)

[HOMEPAGE](#)

[MISSION](#)

[NEWS](#)

[PROJECTS](#)

[EVENTS](#)

[PEOPLE](#)

[PARTNERS](#)

[CAREERS](#)

[CONTACT](#)

[PUBLICATIONS](#)

[APPLICATIONS](#)

[ONTOLOGIES](#)

[LINKED DATA](#)

[SERVICES](#)

[TOOLS](#)

[SOFTWARE](#)

[DOWNLOAD](#)

Wikipedia: *Semantic web suomeksi*

Google this site

Kim Viljanen

Researcher, M.Sc.
(M.Sc. in Computer Science, [University of Helsinki](#))

phone: +358 (0)40 5414654
(mobile)

email: firstname.lastname@tkk.fi

fax: +358-9-451 3356

room: 2541 @ [TUAS-house](#),
Otaniementie 17, Espoo

postal address: Department of Media
Technology
School of Science and
Technology
Aalto University
P.O. Box 15500
FI00076 Aalto
Finland

My URI: <http://iki.fi/digikim#i>



Latest News

[Muistathan: Yhdistetty avoin tieto perjantaina!](#)

Perjantaina 26.3. aamupäivällä Aalto-yliopistolla järjestettävä...

22.3.2010 11:49 by kimvilja

[Yhdistetty avoin tieto Suomessa, perjantaina 26.3.2010](#)

FinnONTO-hanke järjestää yhdessä valtiovarainministeriön...

1.3.2010 17:51 by eahyvone

[Yhteisöllinen semanttinen web 2.0 -tilaisuuden materiaalit nyt jaossa](#)

Perjantaina 15.1. järjestetyn Yhteisöllinen semanttinen web 2.0...

18.1.2010 2:00 by kimvilja

Latest Publications

Tomi Kauppinen, Panu Paakkari, Eetu Mäkelä, Heini Kuittinen, Jari Väättäinen and Eero Hyvönen: [Geospatio-temporal Semantic Web for Cultural Heritage](#)

Jussi Kurki and Eero Hyvönen:

Current activities

- Building [the Finnish Ontology Library Service ONKI](#).
- Distributed content creation and publishing on the Semantic Web, creating [Linked Data](#)

Kala RDF/Turtle-muodossa

```
<http://www.yso.fi/onto/koko/p35193>  
  rdf:type yso-meta:Concept , afo:AFOConcept ;  
  rdfs:subClassOf <http://www.yso.fi/onto/koko/p48176> ;  
  owl:equivalentClass afo:kala , yso:p907 ;  
  skos:prefLabel "kala" @fi , "fish" @en , "fisk" @sv ;  
  skos:related <http://www.yso.fi/onto/koko/p48097> ,  
  <http://www.yso.fi/onto/koko/p35463> ,  
  <http://www.yso.fi/onto/koko/p738> ,  
  <http://www.yso.fi/onto/koko/p4908> ,  
  <http://www.yso.fi/onto/koko/p45714> ,  
  <http://www.yso.fi/onto/koko/p58> ;  
  afo:cc afo:C60_elintarviketieteet , afo:C80_ymparistonsuojelu ;  
  om:definedConcept ysa:Y1674 .
```

2. URI:en muotoilu

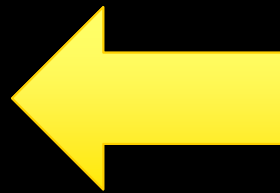
- pysyviä (tai pitkäikäisiä)
- HTTP-muotoisia
- jotka palauttavat lisätietoa

Esimerkki: "kala"

<http://www.yso.fi/onto/koko/p35193>

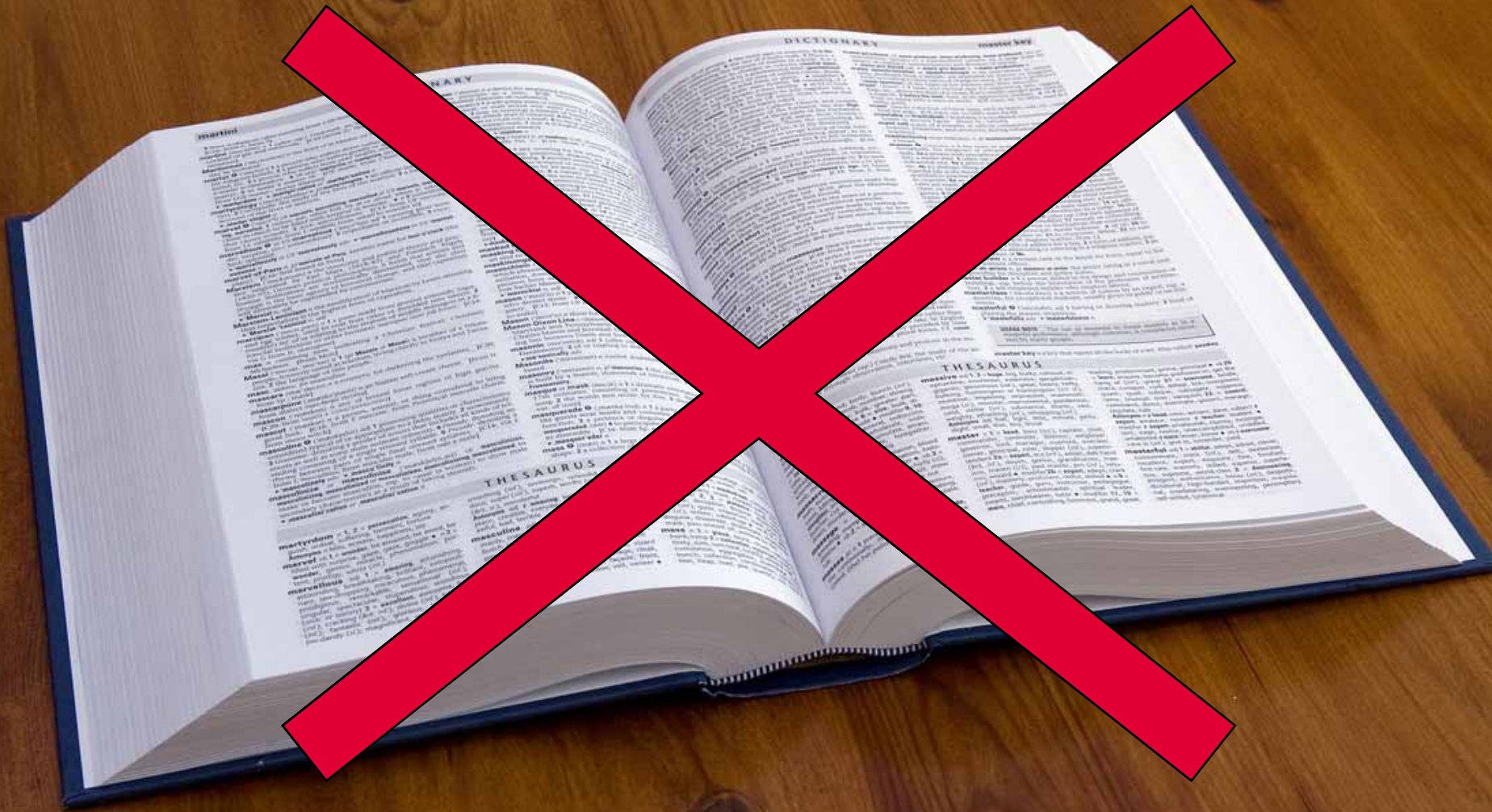
3. skeemat/sanastot

- hyödynnä olemassa olevia suosituksia, esimerkiksi: Dublin core –metatietomalli
- jos luot uusia: suhteuta ne olemassa oleviin



Mitä tietoja tähän liittyy?

Mistä löytyy ”kala”?



DICTIONARY

THESAURUS

THESAURUS

masochism n. 1. 2. a sensation, esp. an
of pain, that is derived from the infliction of
self-harm. 3. a person who derives pleasure
from inflicting pain on himself. 4. a person
who derives pleasure from inflicting pain on
others. 5. a person who derives pleasure
from inflicting pain on others.

masculine

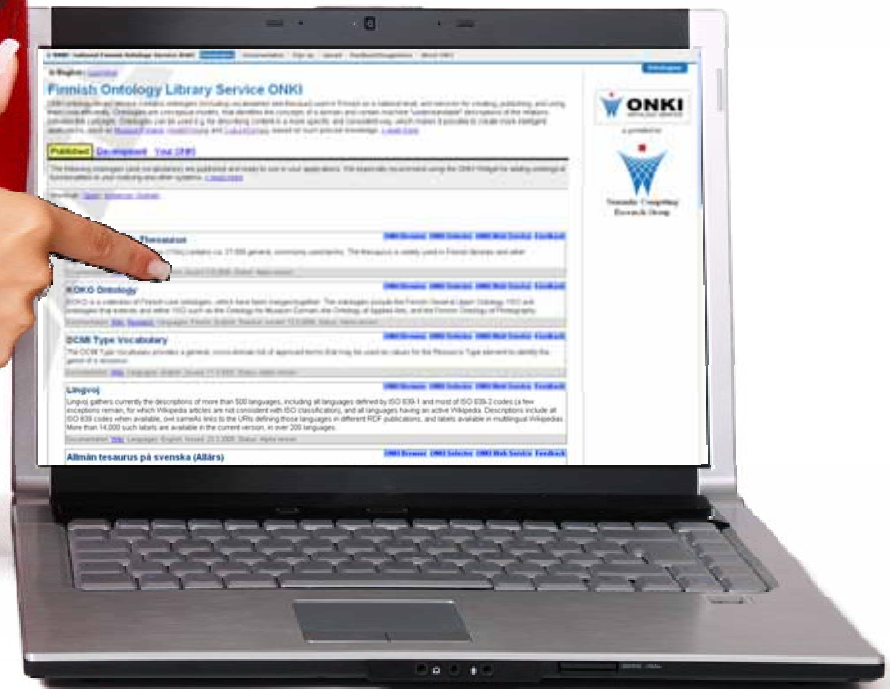
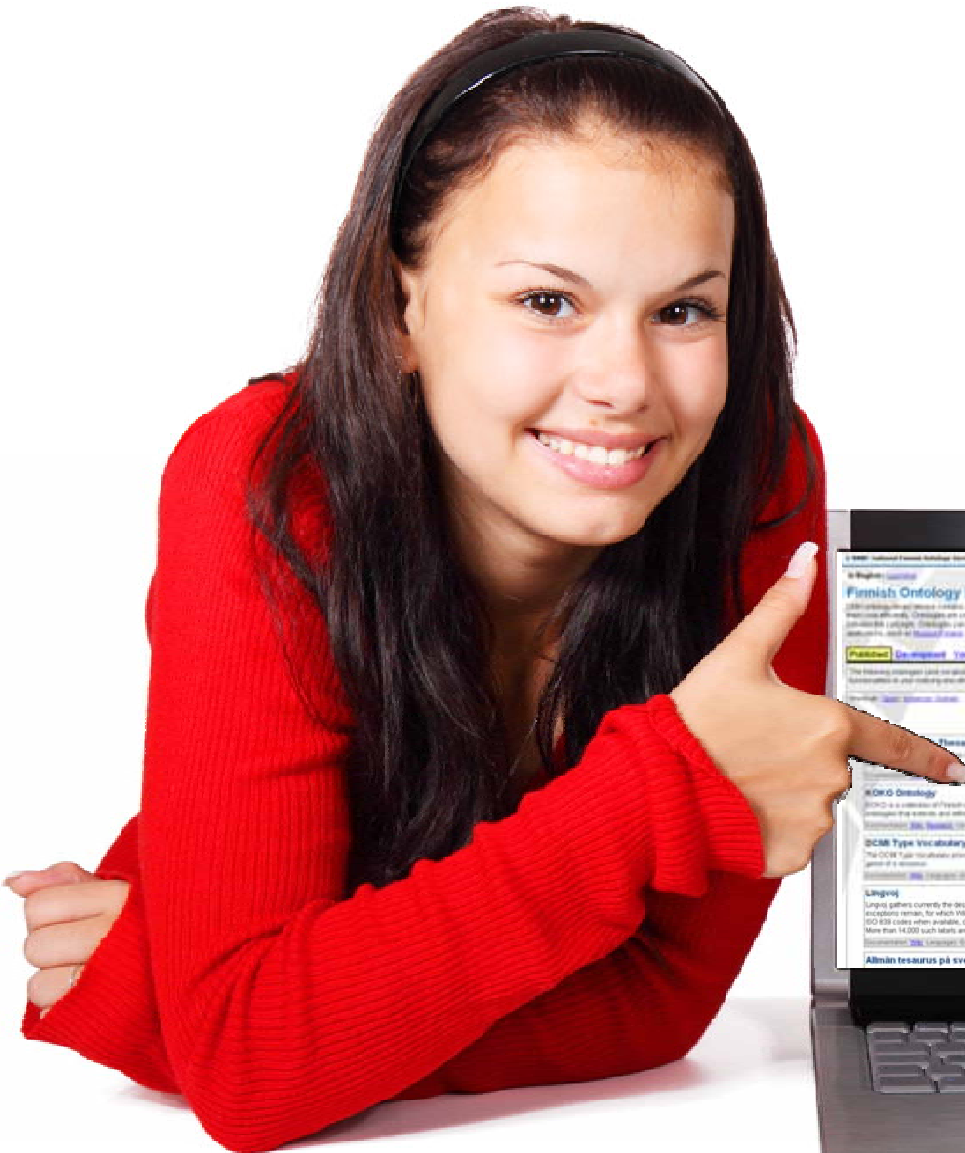
masochist n. 1. a person who derives
pleasure from inflicting pain on himself.
2. a person who derives pleasure from
inflicting pain on others.

masochism n. 1. a sensation, esp. an
of pain, that is derived from the infliction of
self-harm. 2. a person who derives pleasure
from inflicting pain on himself. 3. a person
who derives pleasure from inflicting pain on
others. 4. a person who derives pleasure
from inflicting pain on others.

ONKI

- aina ajantasalla

www.yso.fi



4. URI:n paluuarvo

Mitä osoitteesta

”<http://www.kalatehdas.fi/purkki12345>” saa?



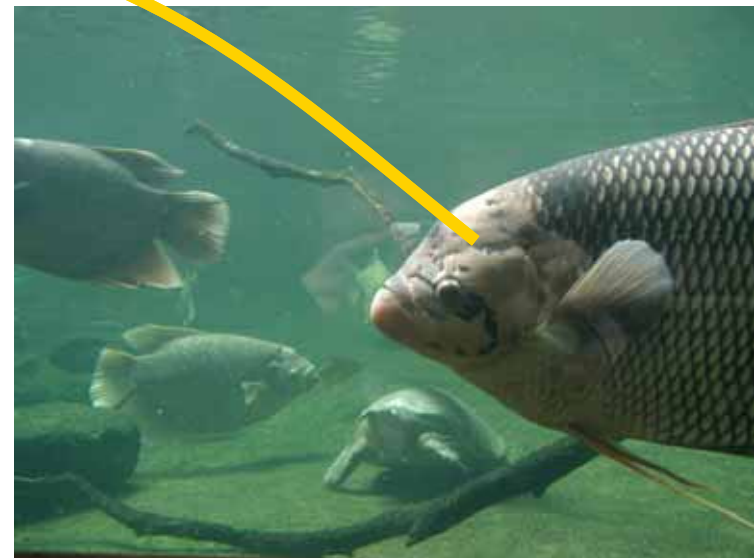
sisältää

<http://www.kalatehdas.fi/purkki12345>

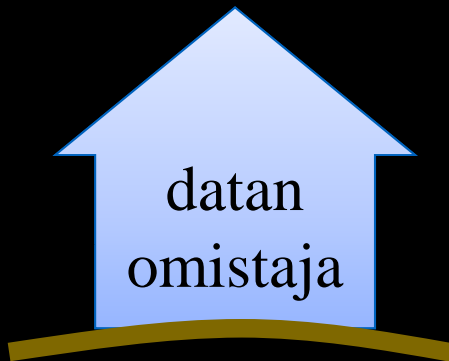
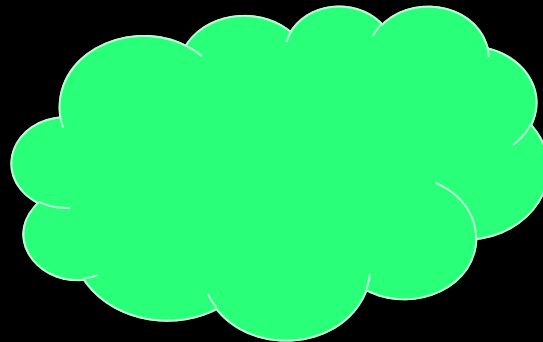
<http://www.yso.fi/onto/koko/p58>

eräänlainen

<http://www.yso.fi/onto/koko/p35193>



5. linkitys



6. julkaiseminen

- RDF-tiedosto
- SPARQL

7. testaus

(=miltä datasi näyttää RDF-
selaimilla)



between **examples**

(1)

(2)

Filter by: relations: (8/8)

length class link conne...

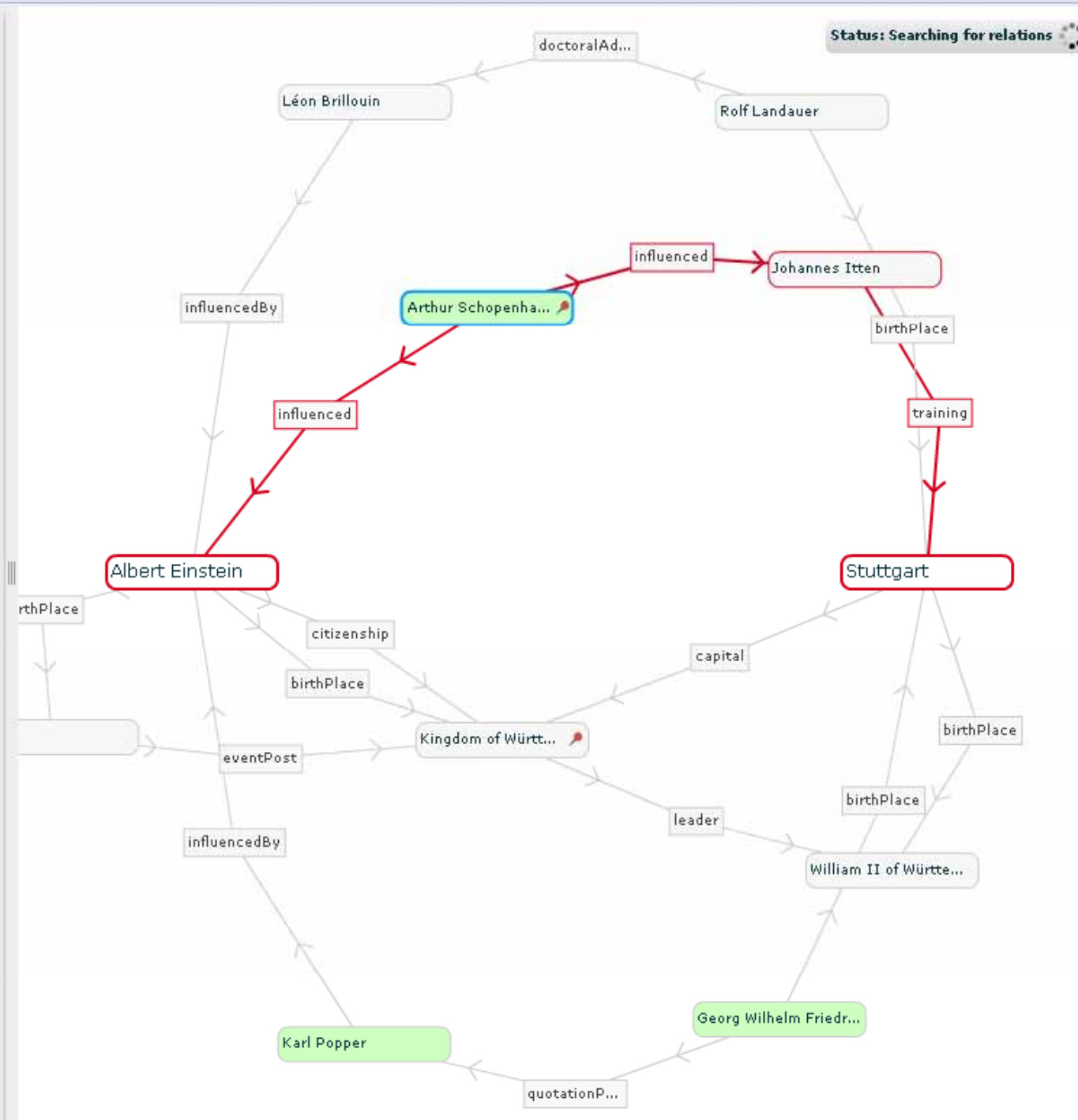
number of objects	num	vi
1	2/2	
2	6/6	

Arthur Schopenha... en ▼

More Infos: dbpedia.org
en.wikipedia.org



Arthur Schopenhauer (22 February 1788 – 21 September 1860) was a German philosopher known for his atheistic pessimism and philosophical clarity. At age 25, he published his doctoral dissertation, *On the Fourfold Root of the Principle of Sufficient Reason*, which examined the fundamental question of whether reason alone can



8. hei maailma!

- miten mainostat omaa dataasi?

Tärkein tapa mainostaa:

sisääntulevat linkit!

(vrt. nykyinen web)



TaskForces/CommunityProjects/LinkingOpenData/DataSets

SWEO Community Project: Linking Open Data on the Semantic Web

navigation

- [Main Page](#)
- [Browse categories](#)
- [Recent changes](#)
- [Help](#)

search

toolbox

- [What links here](#)
- [Related changes](#)
- [Upload file](#)
- [Special pages](#)
- [Printable version](#)
- [Permanent link](#)

Data sets

This page collects RDF data sets that are part of the emerging Web of Linked Data.

Note: a related page exists for [Tabular collation of RDF data set archives](#) that is more amenable to discovery and loading of these data sets into RDF data stores (Quad or Triple stores).

To be part of the Linked Data Web, data has to be accessible as RDF over the HTTP protocol through at least one of the access methods listed below. [The more methods the better](#) (but [avoid aliases](#)). See also tutorial on [How to publish Linked Data on the Web](#).

The page is part of the [SWEO Interest Group Community Projects](#) effort. See also:

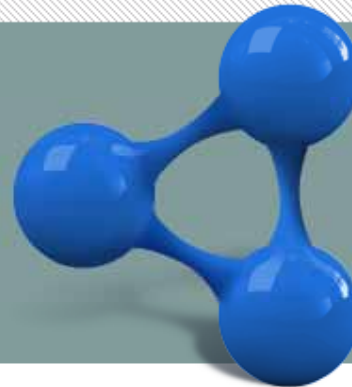
- Richard Cyganiak's click-able version of the [LOD Cloud](#) showing instance relationships among data sets which are linked, directly or indirectly, to DBpedia.
 - An [alternative clickable illustration of the DBpedia Cloud](#), presented by [OpenLink](#), shows the interconnections (and gaps) somewhat more clearly.
- The [UMBEL](#) project provides a clickable diagram of the [UMBEL LOD Constellation](#), an illustration of class linkages across external shared ontologies which are linked, directly or indirectly, to UMBEL.
- [OpenLink's](#) Virtuoso Sponger supports real-time RDF conversion ("RDFization") for this [Dynamic Linked Data Cloud](#), including linkages to the [DBpedia Linked Data Cloud](#), the [UMBEL Class Constellation](#), the [Bio2RDF Linked Data Cloud](#), and other data sets and data dictionaries.
- The [Bio2RDF Linked Data Cloud](#) is described [below](#).
- [Sindice's dataset by topic map](#)
- [The Comprehensive Knowledge Archive Network \(CKAN\)](#)- a registry of open knowledge 'packages', including plenty of open data.

How big is this Web of Linked Data?

- [Statistics on Data sets that are available as Linked Data](#)

Unlocking innovation

Working with UK Public
Sector information and data



Advised by Sir Tim Berners-Lee and Professor Nigel Shadbolt and others, government is opening up data for reuse. This site seeks to give a way into the wealth of government data and is under constant development. We want to work with you to make it better.

We're very aware that there are more people like you outside of government who have the skills and abilities to make wonderful things out of public data. These are our first steps in building a collaborative relationship with you.

Latest news:

- Read about our latest [site changes](#)
- find out how the data.gov.uk team has been [getting involved with the community](#)
- listen to a [Podcast on setting up data.gov.uk](#)

Subscribe by [RSS](#)



Community
Log in / Sign up

Local Data Panel

What is the Semantic Web?

Combining different data sources has never been easy but the Semantic Web will enable data to be joined easily across boundaries.

[Read more](#)

Search Data

Enter keyword(s)

e.g. [education](#), [NHS](#), [crime](#), [transport](#), [environment](#)

[Search](#)

Browse for Data

[List all datasets](#)

[By Public Body](#)

[Common tags](#)

Digital Engagement
Twitter stream

http://seco.tkk.fi/saha3/u2b8f9cf1-6ba1-4285-87d0-7903ee1aaae3

name

message

dataset: MAMO - Maailman nisäkkäät -ontologia

[\[view\]](#) | [\[rdf\]](#) | [\[config\]](#) | [\[remove\]](#)

dateModified	<p>add new literal</p> <input type="text"/> <input type="button" value="add"/>
	<p>[remove] 2010-03-25T14:06:44.727Z</p>
title	<p>add new literal</p> <input type="text"/> <input type="button" value="add"/>
	<p>[remove] MAMO - Maailman nisäkkäät -ontologia</p>
subject	<p>select reference</p> <input type="text"/>
	<p>[remove] biologia [edit] [remove] nisäkkäät [edit]</p>
description	<p>add new literal</p> <input type="text"/> <input type="button" value="add"/>
	<p>[remove] Maailman nisäkkäät -ontologia perustuu Luonnontieteellisen keskusmuseon julkaisemaan Maailman nisäkkäiden suomenkielisten nimien luetteloon. Nimiluettelossa esitetään maailman nisäkkäiden laho-, heimo-, suku- ja lajinimet, ensin tieteelliset nimet, auktori (tieteellisen nimen antaja), englantilainen nimi, vanha suomalainen nimi (jos sellainen on) ja uusi suomalainen nimiehdotus.</p>
homepage	<p>add new literal</p> <input type="text"/> <input type="button" value="add"/>
	<p>[remove] http://www.yso.fi/onki/mamo/</p>
is published under license	<p>select reference</p> <input type="text"/>
example resource	<p>select reference (range unknown)</p>

VoiDTest12

Subject

Biologia	2
Linnut	1
Merenkulku	1
Musiikki	1
Nisäkkäät	1
Paikkatiedot	1

AVIO - Maailman linnut -ontologia (dataset)

Results 5

MAMO - Maailman nisäkkäät -ontologia (dataset)

Merenkulkualan asiasanasto (dataset)

MUSA - Musiikin asiasanasto (dataset)

SUO - Suomalainen paikkaontologia (dataset)

**Kohti yhdistettyä
avointa tietoa...**

Suomen Linked Data -edelläkävijät



KANSALLISKIRJASTO



m-cult



SUOMEN MAATALOUSHUSEO SARKA

Suomen valokuvataiteen museo
Finlands fotografiska museum
The Finnish Museum of Photography



TietoEnator^{TE}



Viestintäteknikka



Finansministeriet
Ministry of Finance



Kartografia ja geoinformatiikka



VALTION TAIDEMUSEO

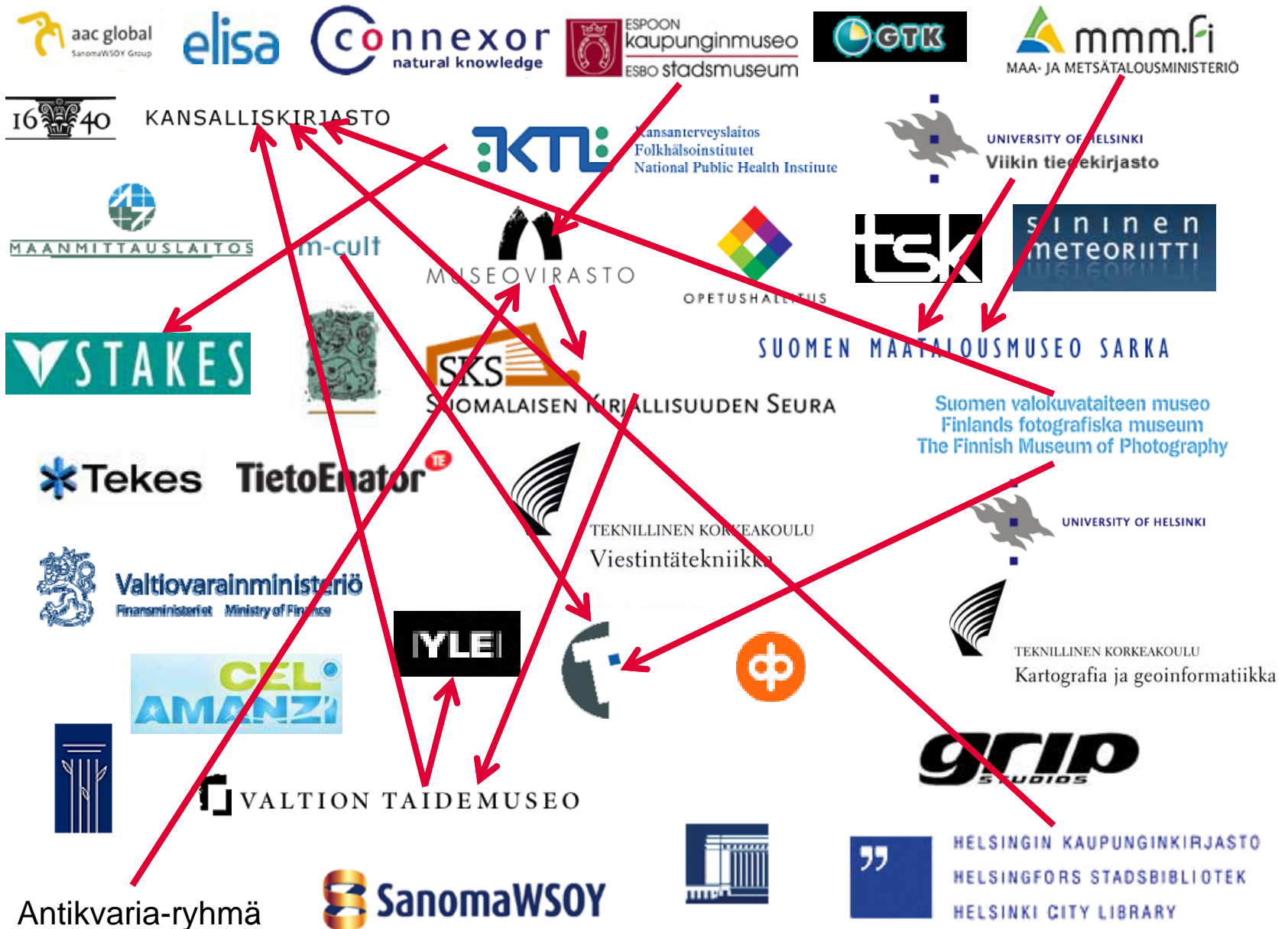


Antikvaria-ryhmä

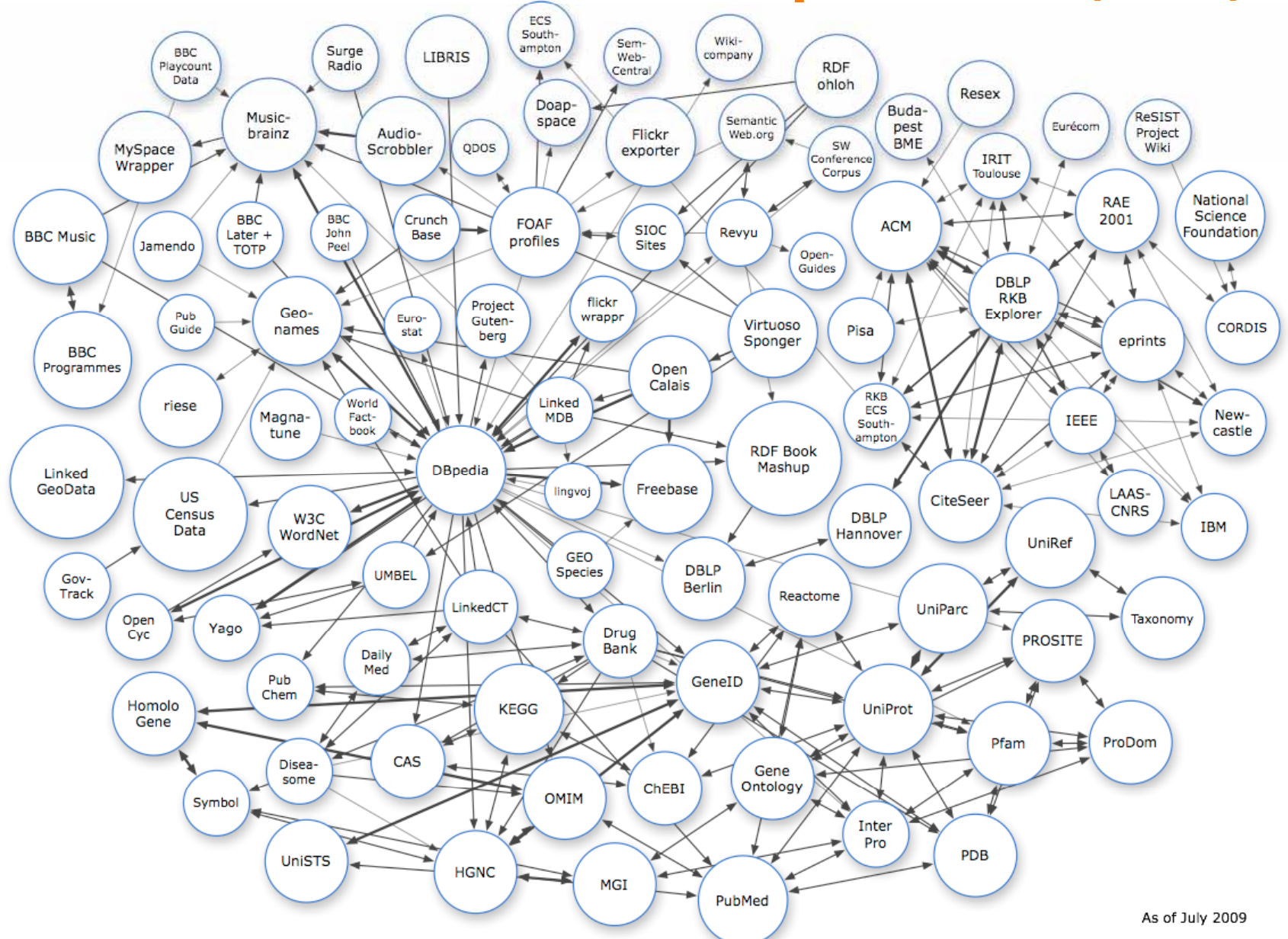


HELSINGFORS STADSBIKLIOTEK
HELSINKI CITY LIBRARY

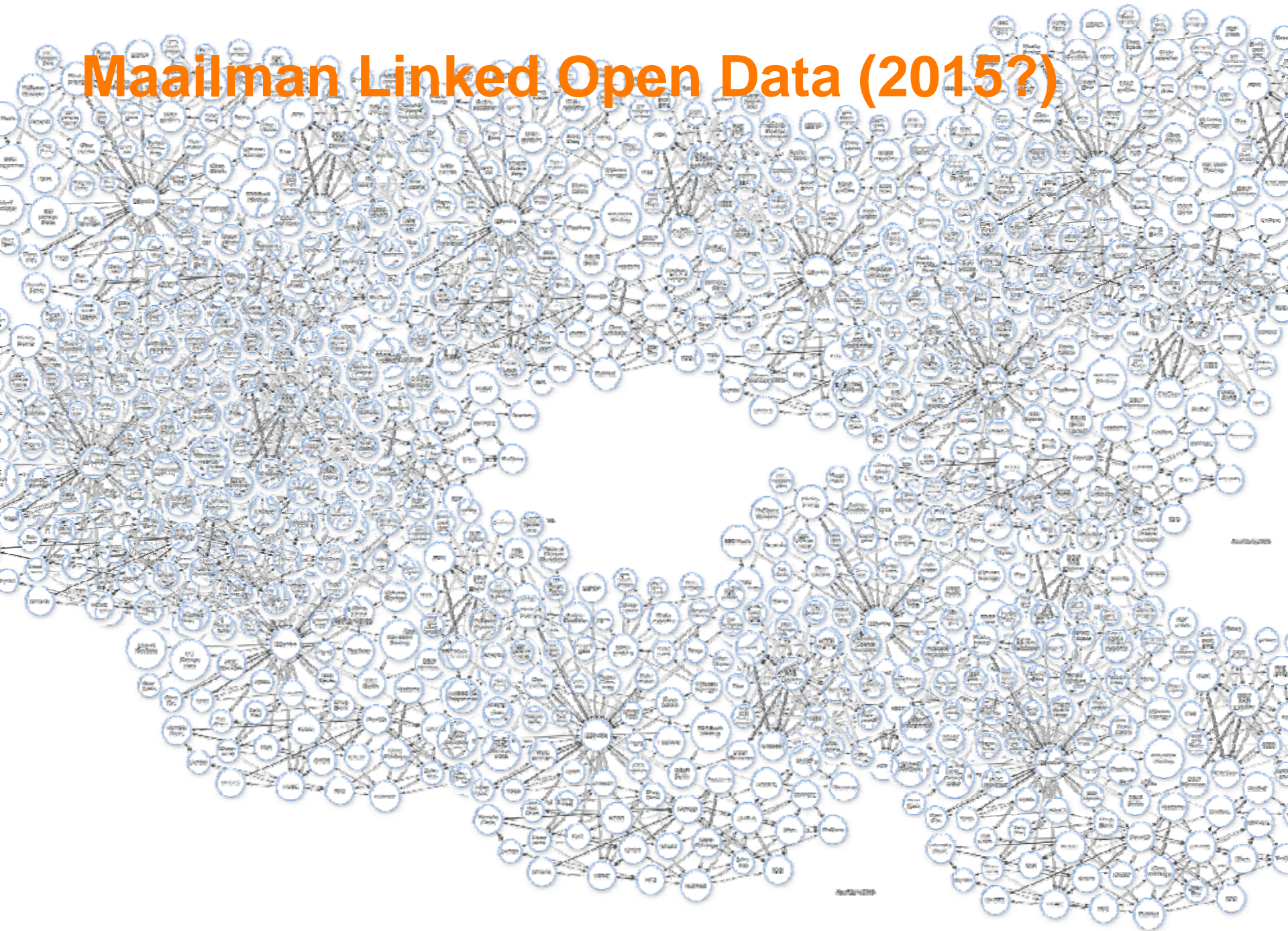
Suomen Linked Data -edelläkävijät



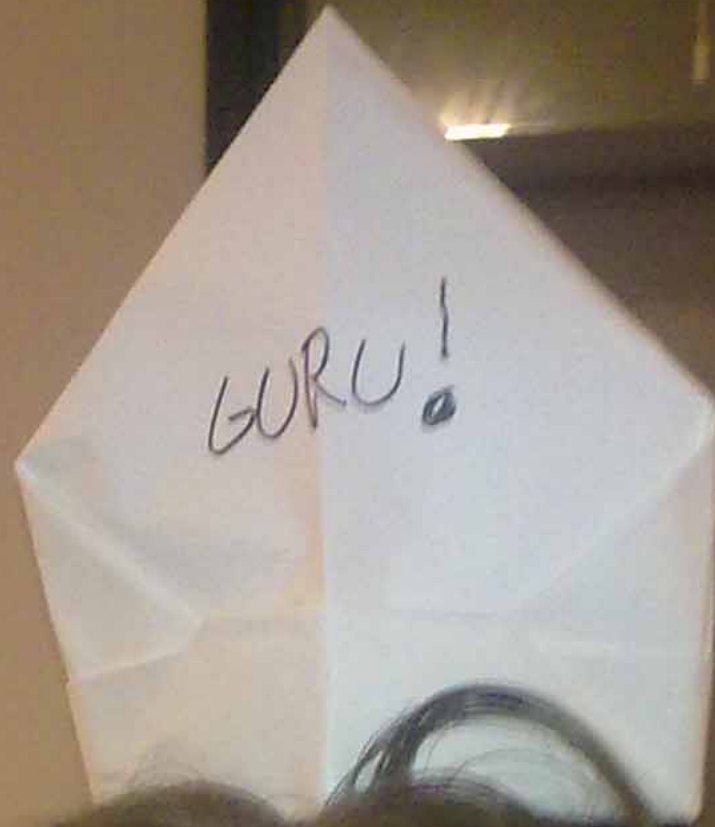
Kansainvälinen Linked Open Data (2009)



Maailman Linked Open Data (2015?)



**...jonka avulla voidaan luoda
ennennäkemättömiä sovelluksia**



Mikä on Linked Open Data Killer App?

Kulttuurisampo: Partamuotia 1800-luvun lopun Suomessa

The screenshot displays a digital archive interface with a grid of items. A pop-up window is open over one item, showing a portrait of a man with a long white beard.

1886: drawing: Isä lukee pikku
1887-1889: painting:
ng: Taiteilijan isän
1891: painting: Juhani Ahon
1892-1894: painting: Gunnar
1893: painting: Herra Ahngerin
1893: painting: Herra Ahngerin muotokuva
1896: photograph:
1897: painting: Ferdinand von
1899: graphic art: Isä
1902: photograp

1893: painting: Herra Ahngerin muotokuva
Sat, 31 Dec 1892 22:00:00 GMT
Sat, 31 Dec 1892 22:00:00 GMT



eri lähteistä kerättyä aineistoa Kulttuurisammossa

Kiitos!

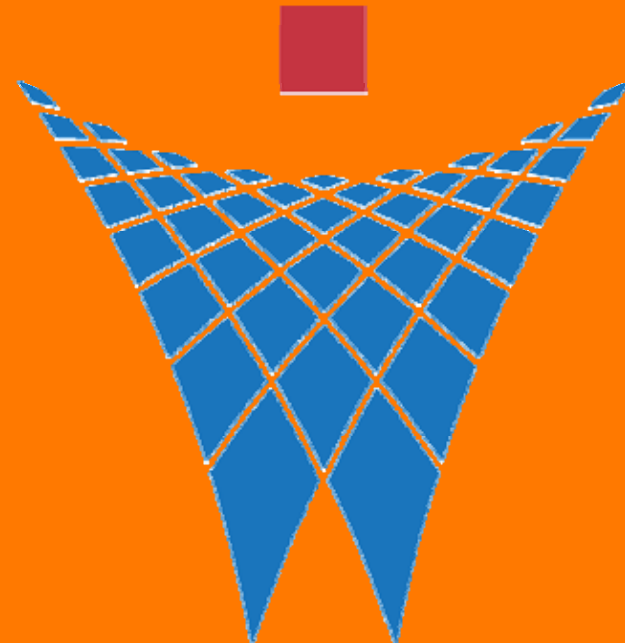
twitter: @secoresearch
<http://www.seco.tkk.fi>

Kim Viljanen

Semanttisen laskennan tutkimusryhmä SeCo

Aalto-yliopisto

kim.viljanen@tkk.fi



Kiitokset/lähteet

- Matias Frosterus
- Eetu Mäkelä
- Osma Suominen

- Sandro Hawke: Linked Open Data in 10 minutes (kalvot)
- Bizer et al: How to Publish Linked Data on the Web, 2007.
- <http://linkeddata.org>
- <http://relfinder.semanticweb.org>
- <http://www.w3.org/DesignIssues/LinkedData.html>
- <http://data.gov.uk>
- <http://www.yso.fi>