



Aalto University  
School of Science



# Semantic Web An Introduction



**CS-E4410 Semantic Web, 13.1.2021**

*Eero Hyvönen*

*Aalto University, Semantic Computing Research Group (SeCo) <http://seco.cs.aalto.fi>*

*University of Helsinki, HELDIG*

*<http://heldig.fi>*

*[eero.hyvonen@aalto.fi](mailto:eero.hyvonen@aalto.fi)*

# Outline

## The idea of Semantic web

## Core technological basis of Semantic web

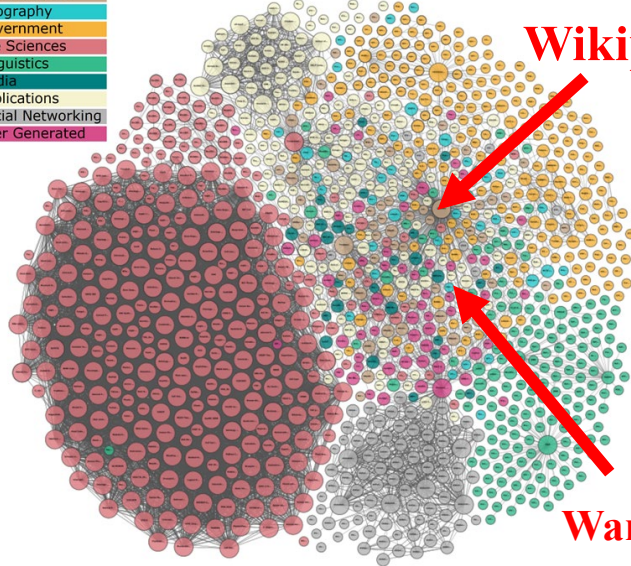
- Metadata, ontologies, reasoning
- Review of the technological solutions and standards

## Application domains



# 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 (2020):  
55 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 Open 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  **FAIR**
  - The value of data increases!
- Creating more intelligent applications for the public, curators, and researchers
  - The machine “understands” linked data!



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



# Fundamental barrier for the development of the Web of Data: machine-”understandability”

The web contents are created for human readers

- HTML, PDF, JPEG, ...

Machine mediates and displays, but does not “understand” contents of the web

- E.g., a Finnish text article



A web service  $\approx$  machine helps human

- Requires machine-”understandability” of the contents

➔ A fundamental contradiction

# How can we build a more intelligent Web?

## 1. Applications are programmed to be more intelligent

- The contents stay as they are
- The machines operate more human-like (Artificial Intelligence)

## 2. Contents are represented in a more intelligent way

- The contents are easier to understand
- Machines stay more or less as they are

## In practice, both ways are needed

- More intelligent systems process more intelligently represented contents

# Approach 1: Develop more intelligent applications using AI

## Automatic interpretation of natural language is difficult

- Free form of the documents
- Semantics of the content

## Non-textual contents

- Pictures, sound, music, video, software, ...
- How to interpret algorithmically unstructured data?

## More than the document itself is needed for interpretation

- Context + common sense needed
- Fundamental problems of Artificial Intelligence, easy for humans!
- Great scientific and technological challenges



# Approach 2: Contents represented in a more intelligent way

## The foundation of Semantic web

- The information is stored in a way that a machine understands it!
- Human helps the machine
  - *Machine can also help in this (user-friendly tools for semantic content creation)*

## The development was boosted in the beginning of the 2000s

- W3C Semantic Web Activity 2001
- W3C Web Services Activity 2002

# Web Generations

## 1G WWW: early 1990's

- WWW pages for human interpretation
- HTML language

## 2G WWW: late 1990's

- Structured web documents for human/machine interpretation
- XML-based languages

## 3G WWW: Semantic Web, Web of Data, 2000's

- Explicit meaning of documents for human/machine use
- RDF-based languages

⇒ **Semantics = new foundation for intelligent web services**

- Semantic = “understandable” to machines

# Limitations of non-semantic web: case MuseumFinland

```
<artifact>  
  <id>NBA:H26069:467</id>  
  <target>cup and plate</target>  
  <material>porcelain</material>  
  <creationLocation>Germany</creationLocation>  
  <creator>Meissen</creator>  
</artifact>
```



- This metadata cannot answer the following questions:
  - Find all vessels?
  - Find all ceramic products?
  - Find artifacts manufactured in Europe?
  - Does the city of Meissen manufacture ceramics?

# Semantic Web solution: ontologies

NBA-H26069-467

:object "cup and plate" ;

:object\_concept **object:cup** ;

:object\_concept **object:plate** ;

:material "porcelain" ;

:material\_concept **object:porcelain** ;

:creationPlace "Germany" ;

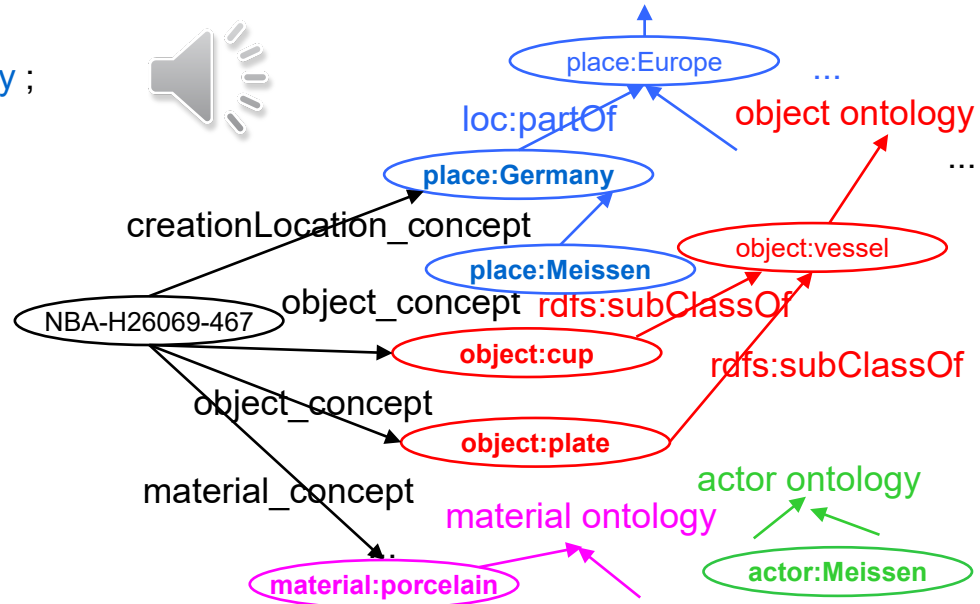
:creationPlace\_concept **place:Germany** ;

:creator "Meissen"

:creator\_concept **actor:Meissen** .



place ontology



Find all vessels?  
Find all ceramic products?  
Find artifacts manufactured in Europe?  
Does the city of Meissen manufacture ceramics?

# Case Rijksmuseum Amsterdam: CHIP Demonstrator

## Example in Turtle notation

- VRA metadata schema  
(extension of Dublin Core)
- (Aroyo et al., 2007)

```
rijks:artefactSK-C-K
  vra:type vra:Work ;
  vra:title "The Night Watch" ;
  vra:date "1642" ;
  vra:creator: 500011051 ;      # Rembrandt
  vra:subject iconclass:45F31 ; # Call to arms
  vra:culture tgn:7006952 ;    # Amsterdam
  vra:material aat:30015050 .  # Oil paint
```



A resource in the TGN  
ontology / vocabulary

# Amsterdam in TGN



This screenshot shows the 'TGN Full Record Display' for 'Amsterdam (inhabited place)'. The browser address bar shows the URL: [http://www.getty.edu/vow/TGNFullDisplay?find=Amsterdam&place=&nation=&prev\\_page=1&english=Y&](http://www.getty.edu/vow/TGNFullDisplay?find=Amsterdam&place=&nation=&prev_page=1&english=Y&). The page provides detailed information about the place, including its coordinates, historical notes, names in various languages, hierarchical position, place types, sources, and subjects.

**Amsterdam (inhabited place)**

Coordinates: Lat: 52 21 00 N degrees minutes Long: 004 54 00 E degrees minutes  
Lat: 52.3500 decimal degrees Long: 4.9000 decimal degrees

**Notes** Located on over 90 islands in the IJ arm of the IJsselmeer. Early inhabitants built dikes on both sides of the Amstel River to prevent flooding, and a dam was built between the dikes in 1270. Chartered in 1306. Became affluent in the 15th century due to trade with Baltic seaports, and was the financial center of the world by 17th century. United Dutch East India Company was founded in 1602, followed by the West India Company in 1621. Capital of the Batavian Republic under Napoleon, later of the kingdom of Holland, and became part of the French Empire in 1810. Under German occupation from 1940-1945. Center of the world's diamond trade.

**Names**

**Amsterdam**  
(preferred, C, V, N, English-P, Dutch-P)

**Amstel-dam** (N, V, N) ..... documented in 13th cen., meaning 'dam on the Amstel' [over]

**Amsteldam** (N, V, N)

**Amsteldamme** (N, V, N) ..... earliest form of the name, 13th cen.

**Amsteddamum** (N, V, N)

**Amsterdamum** (N, V, N)

**Amstredamum** (N, V, N)

**Amstredamense oppidum** (N, V, N)

**Hierarchical Position:**

- World (face)
- Europe (continent)
- Netherlands (nation)
- North Holland (province)
- Amsterdam (inhabited place)

**Place Types**

**inhabited place** (preferred, C) ..... there possibly was a Roman settlement in the area; modern town probably originated as a fishing village in 13th century

**city** (C)

**capital** (C) ..... nominal capital of The Netherlands, though government is located in s-Gravenhage

**port** (C)

**Sources and Contributors:**

**Amsterdamum** ..... [VP] Orbis Latinus (1971) 18

**Amstel-dam** ..... [VP] Encyclopaedia Britannica (1988) 1, 357

**Amsterdam** ..... [VP] Knopf Guides, Amsterdam (1993) 28

**Amsteldamme** ..... [VP] Knopf Guides, Amsterdam (1993) 28

**Amsteddamum** ..... [VP] Orbis Latinus (1971) 18

**Amstredamum** ..... [VP] Orbis Latinus (1971) 18

**Amstredamense oppidum** ..... [VP] Orbis Latinus (1971) 18

**Amsterdamum** ..... [VP] Orbis Latinus (1971) 18

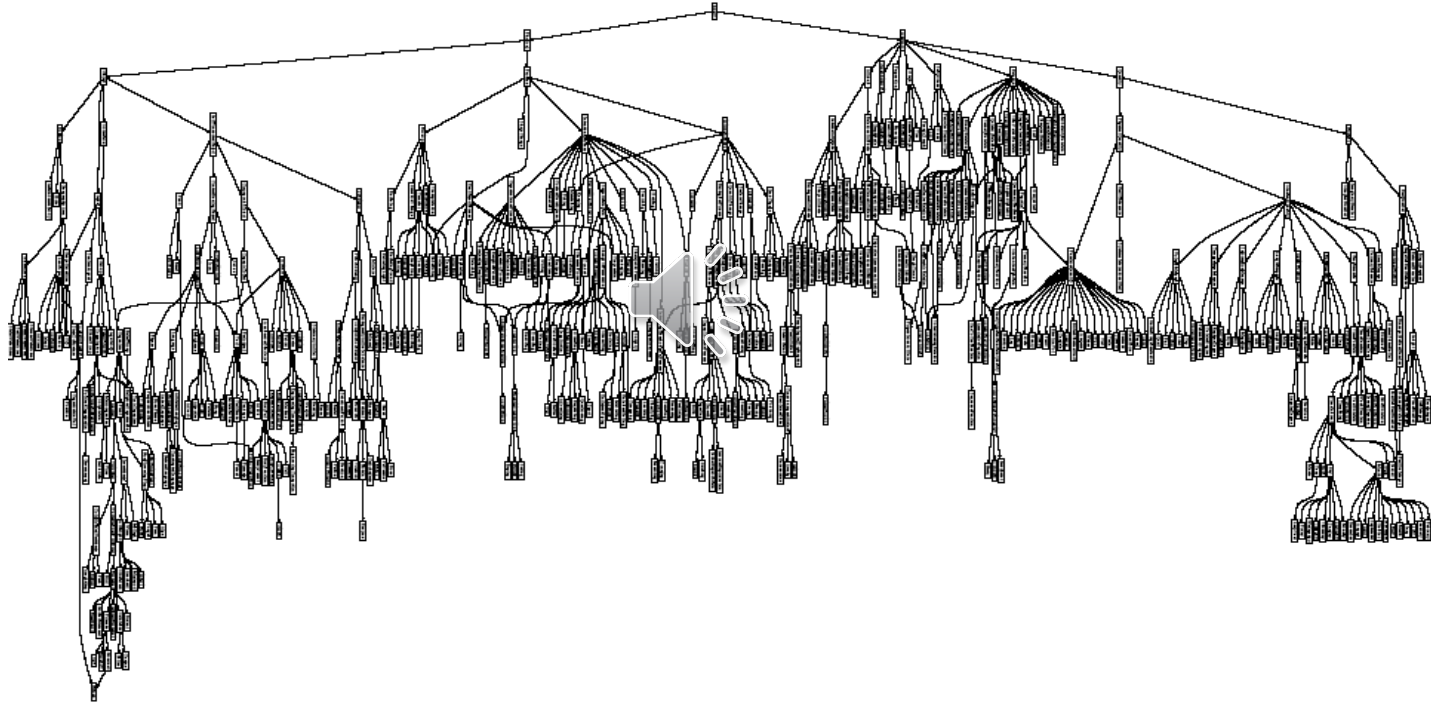
**Amsterdam** ..... [BH, GRIPSC, VP] Canby, Historic Places (1984) 1, 30  
Columbia Lippincott Gazetteer (1961)  
Encyclopaedia Britannica (1988) 1, 357-358  
Encyclopaedia Britannica (1988) 1, 357-358  
Times Atlas of the World (1992) 12  
Webster's Geographical Dictionary (1984)

**Amstredamum** ..... [VP] Orbis Latinus (1971) 18

**Subject:** ..... [BH, GRIPSC, VP] Canby, Historic Places (1984) 1, 30  
Columbia Lippincott Gazetteer (1961)  
Encyclopaedia Britannica (1988) 1, 357-358  
Knopf Guides, Amsterdam (1993) 28  
Orbis Latinus (1971) 18  
Phadon Art Guide: Holland (1987) 18  
Times Atlas of the World (1992) 12  
Webster's Geographical Dictionary (1984)  
Webster's Geographical Dictionary (1988) 46

**Notes** ..... [VP]

# An Ontology Concept Hierarchy: Standard Upper Merged Ontology SUMO

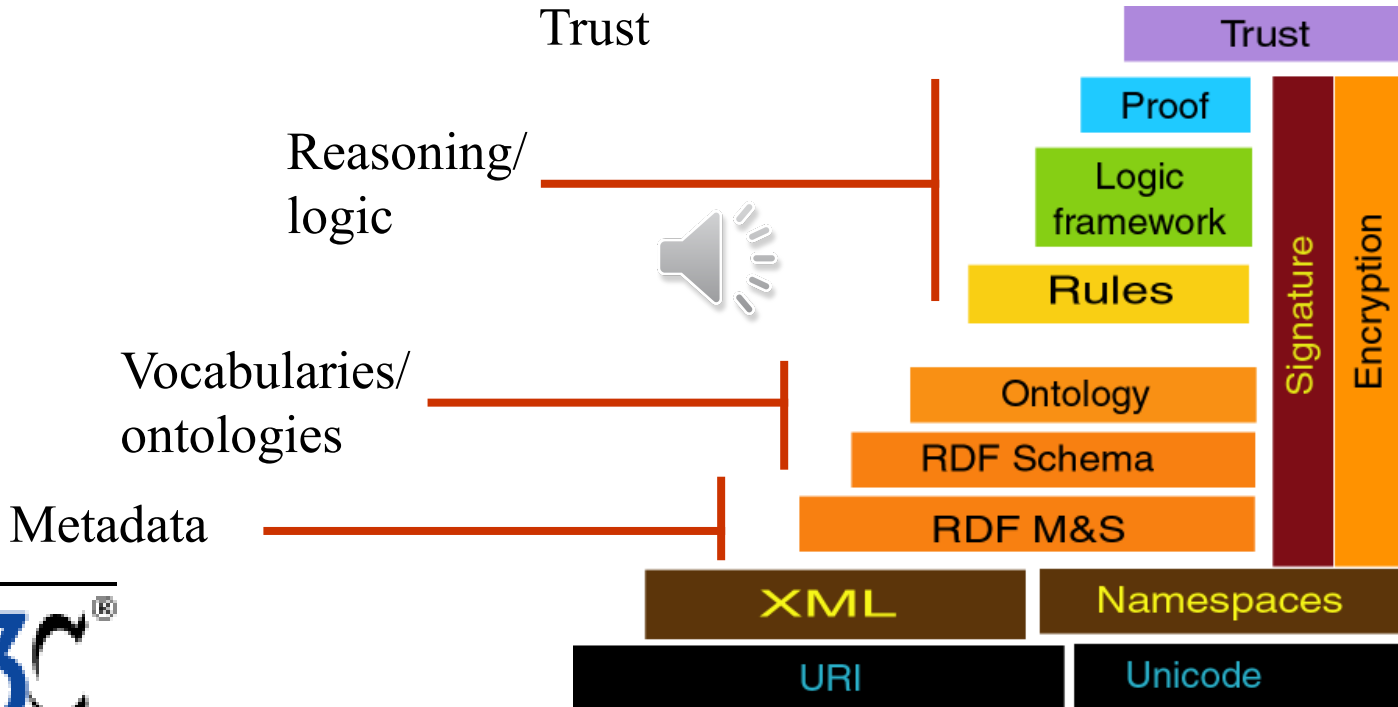


# Technological basis of Semantic Web



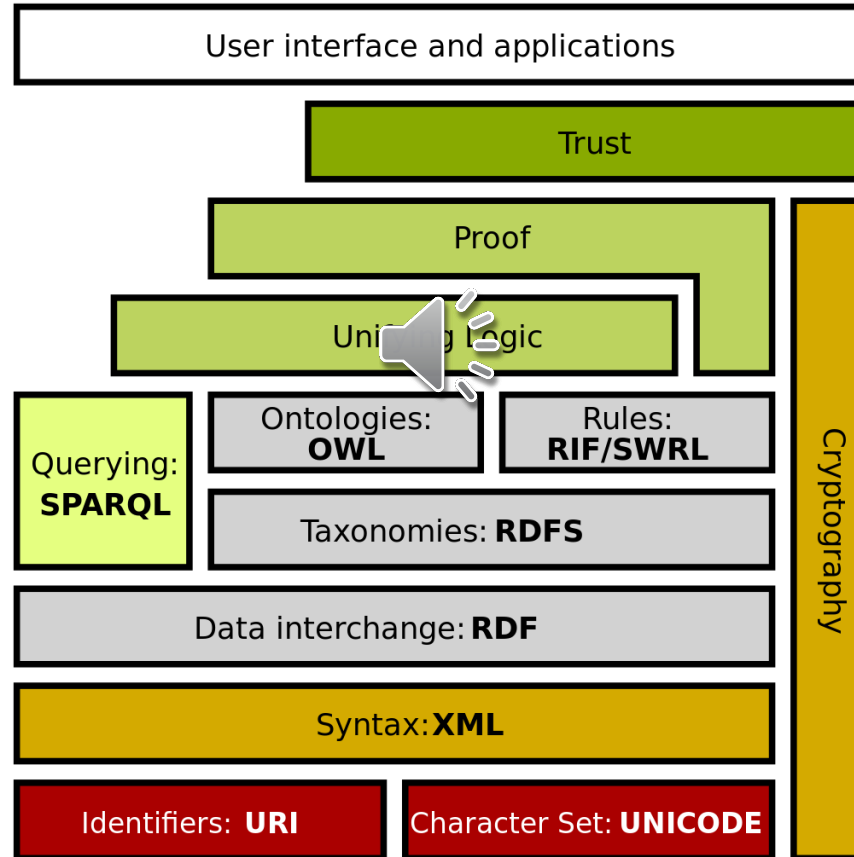


# The Original "Layer Cake Model" of Semantic Web



(Tim Berners-Lee)

# Newer Version of Layer Cake Model




# Metadata Level





# The Semantic Web solution for Metadata: RDF Resource Description Framework

- General metadata description model and language for web resources
- Relational model, *not* a syntax (as opposed to XML)
  - *RDF description = directed (knowledge) graph*
- Semantics is defined based on logic
- RDF has syntaxes/serializations,  *to*
  - *XML-based RDF/XML, especially for machines*
  - *Simpler notations (Turtle, N-triples, N3) for humans*
- Standardized and commonly used
  - *W3C draft 1999*
  - *W3C recommendation RDF 1.0, 10.2.2004*
  - *W3C recommendation RDF 1.1, 25.2.2014*



# RDF Example

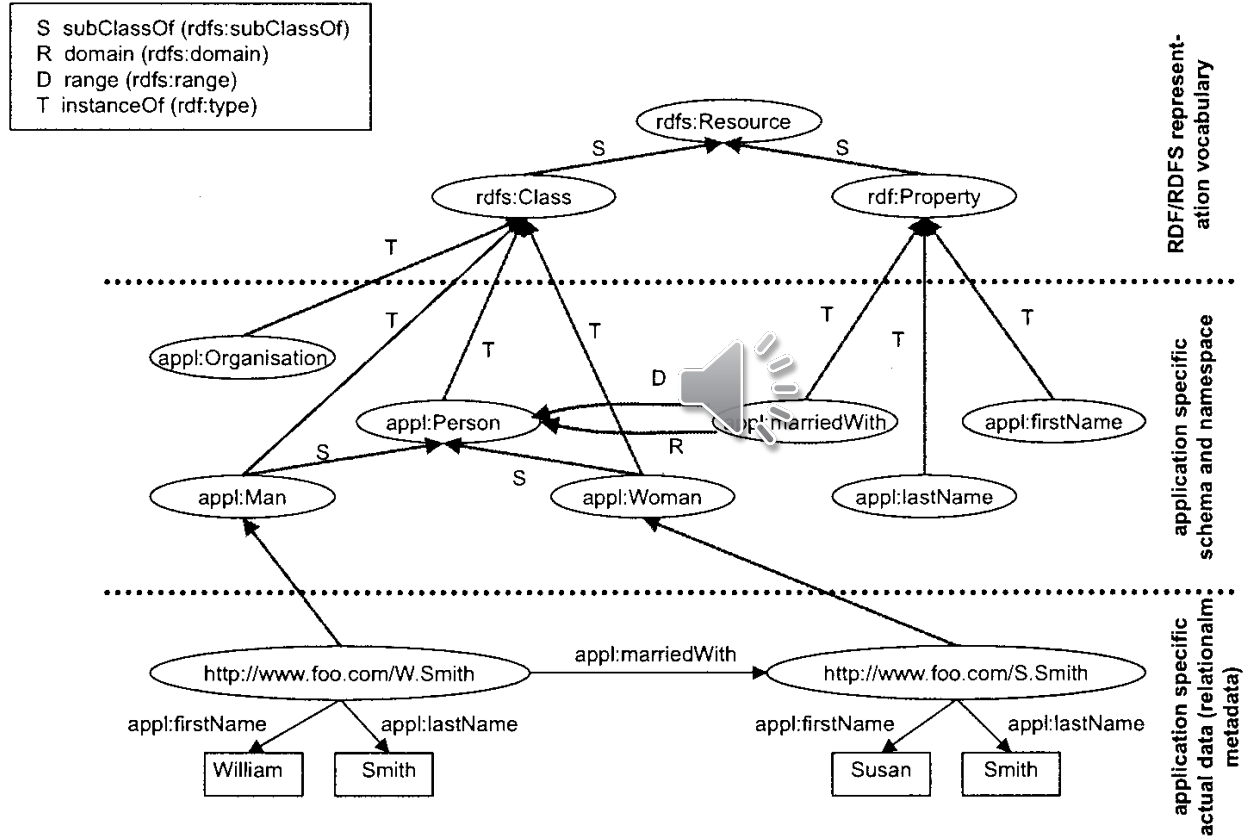


Figure 3.4. An RDF-Schema Example

(Maedche, 2002)

# Metadata Schemas

## Standardized templates for representing (meta)data

- A set of elements (properties) describing object types
  - *E.g author, publisher, and publishing year of books*
- Value specifications for the properties describing individual objects
  - *Literal values for data (text, number, date, ...)*
  - *URI identifier for related concepts/resources*

## Different content types typically require different schemas (element sets)

- E.g. books, persons, paintings, places, ...

# Example: Dublin Core metadata schema for describing web documents

## Set of 15 general properties for different content types

- Dublin Core Metadata Element Set (ISO Standard 15836)
  - *Title*
  - *Creator*
  - *Subject*
  - *Description*
  - *Publisher*
  - *Contributor*
  - *Data*
  - *Type*
  - *Format*
  - *Identifier*
  - *Relation*
  - *Source*
  - *Language*
  - *Coverage*
  - *Rights*





# Metadata Schema in HealthFinland

**Table 1.** HEALTHFINLAND Metadata Schema. Obligatory fields are marked in **bold**. Cardinalities are presented in the column C.

	Name	QName	C	Value type	Value range
General metadata	<b>Identifier</b>	<b>dc:identifier</b>	1	URI	
	Locator	ts:url	0..1	URL	
	<b>Title</b>	<b>dc:title</b>	1 <sup>a</sup>	Free text	Non-empty string.
	<b>Abstract</b>	<b>dcterms:abstract</b>	1 <sup>a</sup>	Free text	Non-empty string.
	<b>Language</b>	<b>dc:language</b>	1..*	String	RFC 3066
	<b>Publication time</b>	<b>dcterms:issued</b>	1	String	W3CDTF (ISO 8601)
	Acceptance time	dcterms:dateAccepted	0..*	String	W3CDTF (ISO 8601)
	Modification time	dcterms:modified	0..*	String	W3CDTF (ISO 8601)
	<b>Publisher</b>	<b>dc:publisher</b>	1..*	Instance	foaf:Organization
	Creator	dc:creator	0..*	Instance	foaf:Organization, foaf:Person or foaf:Group
Content classification	<b>Subject</b>	<b>dc:subject</b>	1..*	Concept	ICD, MeSH and HPMulti Ontologies
	<b>Audience</b>	<b>dcterms:audience</b>	1..*	Concept	Audience Ontology
	<b>Genre</b>	<b>ts:genre</b>	1..*	Concept	Genre Ontology
	<b>Presentation type</b>	<b>dc:type</b>	1..*	Concept	DCMI Type vocabulary
	<b>Format</b>	<b>dc:format</b>	1	String	IANA MIME types
	<b>Medium</b>	<b>dcterms:medium</b>	1	Concept	Medium Ontology
	Spatial coverage	dcterms:spatial	0..*	String or concept	DCMI Point, DCMI Box or Location Ontology
	Temporal coverage	dcterms:temporal	0..*	String or concept	W3CDTF, DCMI Period or Time Ontology
Relations	Part of	dcterms:isPartOf	0..*	Document	URI
	Rights	dc:rights	0..*	Free text or document	URI or textual description
	Source	dc:source	0..*	Free text or document	URI (e.g., ISBN) or bibliographical reference
	Reference	dcterms:references	0..*	Free text or document	URI (e.g., ISBN) or bibliographical reference
	Translation of	ts:isTranslationOf	0..*	Document	URI
	Format of	dcterms:isFormatOf	0..*	Document	URI

<sup>a</sup> Multilingual values are allowed, but only one value in each language.

# HealthFinland portal: Maija's eyeglasses – PDF document on the web



The screenshot shows a Windows Internet Explorer browser window displaying a PDF document. The address bar shows the URL: [http://www.ttl.fi/NR/rdonlyres/142A177B-FE02-4A77-A434-3049C37EBC61/0/maijan\\_lasit\\_tulosta.pdf](http://www.ttl.fi/NR/rdonlyres/142A177B-FE02-4A77-A434-3049C37EBC61/0/maijan_lasit_tulosta.pdf). The document content is in Finnish and discusses Maija's new eyeglasses. A large speaker icon is overlaid on the left side of the document content area.

TTL - Maijan uudet näyttöpäätelasit <http://www.ttl.fi/ergonomia>

## Maijan uudet näyttöpäätelasit

Sopivien silmälasien hankinta tietotyöhön voi useinkin olla pulmallista. Seuraavassa näet, miten ikänsäköinen Maija päätyi silmälasiratkaisuunsa.


Maija sai lähinnä huonontuessa aluksi lulasit. Hän hankki kymmenen vuoden aikana kolmet lulasit: uudet aina hiukan entistä vahvemmat.

Päätteitä lukeminen alkoi käydä ongelmalliseksi. Maija ei osannut kirjoittaa sokkona joten näppäimistölle näkeminen oli myös välttämätöntä. Niska ja hartat jomottivat, silmiä kurveli ja vasemmassa silmässä oli elohiiri vuokkoauksia.

Lulasit käytössä.  
Maija joutuu istumaan epämiellyttävään lähellä päätettä. Kynänpäiden voimakas koukistaminen sitoo käsiä ja vaikeuttaa näppäilytyötä.

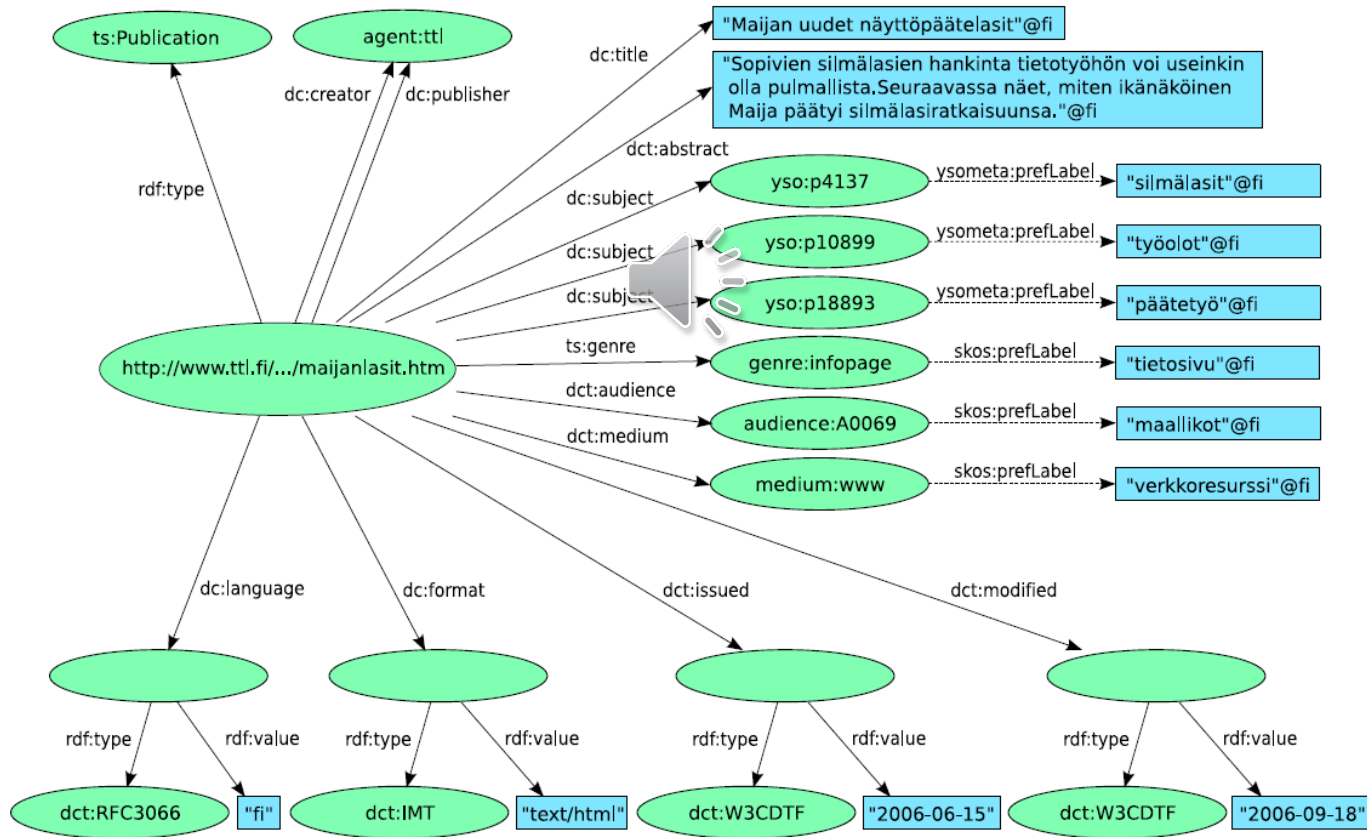


Lulasit käytössä.  
Maija joutuu taivuttamaan niska voimakkaasti eteen ahdakseen näppäimet. Päätteen ja näppäimistön vuorottainen katselu aiheuttaa jatkuvaa niskan liikettä.



1 of 4

# Maija's eyeglasses: metadata in RDF form



# Ontology Level



# What is an ontology?

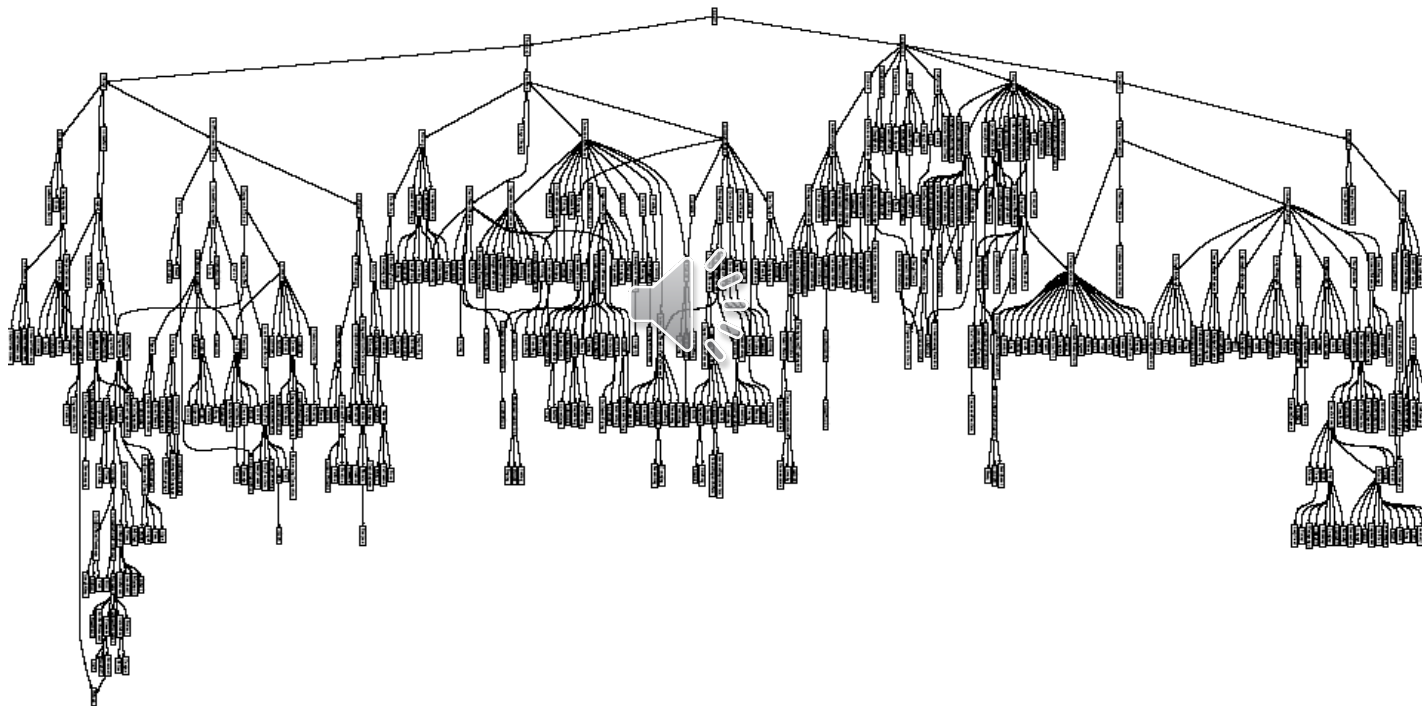
**“An ontology is an explicit specification of a conceptualization  
...definitions need to be couched in some common formalism”**

**(Gruber, 1993)**

- *Explicit: machine can understand*
  - *Formal: precisely defined*
  - *Common (shared): communication is possible*
- **Defines the concepts/objects and their relations in a given application domain**
  - **A first requirement for the humans and machines to understand each other**



# Standard Upper Merged Ontology SUMO



# SUMO principal distinctions

- [Entity](#)
  - [Physical](#)
    - [Object](#)
      - [SelfConnectedObject](#)
        - [Substance](#)
        - [CorpuscularObject](#)
        - [Food](#)
      - [Region](#)
      - [Collection](#)
      - [Agent](#)
    - [Process](#)
  - [Abstract](#)
    - [SetOrClass](#)
    - [Relation](#)
    - [Quantity](#)
      - [Number](#)
      - [PhysicalQuantity](#)
    - [Attribute](#)
    - [Proposition](#)

# SUMO Object:

- [Object](#)
  - [SelfConnectedObject](#)
    - [Substance](#)
      - [PureSubstance](#)
        - [ElementalSubstance](#)
          - [Metal](#)
          - [Atom](#)
          - [SubatomicParticle](#)
            - [AtomicNucleus](#)
            - [Electron](#)
            - [Proton](#)
            - [Neutron](#)
        - [CompoundSubstance](#)
          - [Water](#)
          - [Molecule](#)
      - [Mixture](#)
        - [Solution](#)
        - [Mineral](#)
        - [BodySubstance](#)
      - [BiologicallyActiveSubstance](#)
        - [Nutrient](#)
        - [Hormone](#)
      - [CorpuscularObject](#)
        - [OrganicObject](#)
          - [Organism](#)
          - [AnatomicalStructure](#)
        - [Artifact](#)
        - [ContentBearingObject](#)
      - [Food](#)
    - [Region](#)
      - [GeographicArea](#)
      - [AstronomicalBody](#)
      - [Hole](#)
    - [Collection](#)
      - [Group](#)
        - [GroupOfPeople](#)
        - [Organization](#)
    - [Agent](#)
      - [Organism](#)
      - [Group](#)
      - [GeopoliticalAgent](#)
      - [SentientAgent](#)








[New Search](#)











[Previous Page](#)

[View Selected Records](#)

[Clear All](#)

Click the  icon to view the hierarchy.

Check the boxes to view multiple records at once.

-  Top of the AAT hierarchies
-  .... Associated Concepts Facet
-  ..... Associated Concepts (hierarchy name)
-  .... Physical Attributes Facet
-  ..... Attributes and Properties (hierarchy name)
-  ..... Conditions and Effects (hierarchy name)
-  ..... Design Elements (hierarchy name)
-  ..... Color (hierarchy name)
-  .... Styles and Periods Facet
-  ..... Styles and Periods (hierarchy name)
-  .... Agents Facet
-  ..... People (hierarchy name)
-  ..... Organizations (hierarchy name)
-  ..... Living Organisms (hierarchy name)
- ..... agents (general) [N]
-  .... Activities Facet
-  ..... Disciplines (hierarchy name)
-  ..... Functions (hierarchy name)
-  ..... Events (hierarchy name)
-  ..... Physical and Mental Activities (hierarchy name)
-  ..... Processes and Techniques (hierarchy name)
- ..... activities (general context)
-  .... Materials Facet
-  ..... Materials (hierarchy name)
-  .... Objects Facet
-  ..... Built Environment (hierarchy name)
-  ..... Components (hierarchy name)
-  ..... Furnishings and Equipment (hierarchy name)
-  ..... Object Genres (hierarchy name)
-  ..... Object Groupings and Systems (hierarchy name)
-  ..... Visual and Verbal Communication (hierarchy name)
-  ..... <temporary alphabetical list: objects>
-  ..... <temporary list/DIBAM-CDBP-SNPC test TRP contributions holding>
-  .... Brand Names Facet
-  ..... Brand Names (hierarchy name)



**AAT Art & Architecture Thesaurus**  
 - maintained by Getty Research Institute  
 - 7 main classes, 125 000 concepts

# Universal List of Artist Names ULAN

- Over 300 000 artists with 720 000 names as Linked Open Data (2018)

## Example: Eero Saarinen data



Click the icon to view the hierarchy.

[Semantic View \(JSON, JSONLD, RDF, N3/Turtle, N-Triples\)](#)

ID: 500006141

Page Link: <http://vocab.getty.edu/page/ulan/500006141>

Record Type: [Person](#)

**Saarinen, Eero** (American architect, designer, 1910-1961)

**Note:** Son of Eliel Saarinen and Louise (Loja) Gesellius, the sculptor and weaver. Eero Saarinen emigrated with his family to the United States in 1923. He attended the Académie de la Grande Chaumière, Paris, France, (1929-1930/1931), studied architecture at Yale University, New Haven, Connecticut, and worked in his father's architectural firm, Saarinen and Saarinen, in Ann Arbor, Michigan (1936/1937-1941). He was partner with his father and J. Robert Swanson as Saarinen-Swanson-Saarinen in Ann Arbor (1941-1947) and partner with his father as Saarinen and Associates in Ann Arbor (1947-1950). He directed Eero Saarinen and Associates, Birmingham, Michigan, 1950-1961. He acted as a consultant for the Architects Advisory Panel for the Unesco buildings (built 1955-1958) in Paris, France. American architect.

**Names:**

**Saarinen, Eero** ([preferred](#), [V.index](#), [English-P.NA,U](#))

**Eero Saarinen** ([V.display](#))

סריקן, אירו ([U.Hebrew-P.NA,U](#))

**Nationalities:**

American ([preferred](#))

Finnish

**Roles:**

artist ([preferred](#))

designer

architect

furniture designer

**Gender:** male

**Events:**

active: [United States \(North and Central America\)](#) ([nation](#))

active: [Finland \(Europe\)](#) ([nation](#))

**Related People or Corporate Bodies:**

child of .... [Saarinen, Eliel](#)

..... (Finnish architect, 1873-1950, active in the United States) [500027014]

employee of .... [Eklund, Jari](#)

..... (Finnish architect, 1876-1962) [500069436]

employee was .... [Pelli, Cesar](#)

..... (American architect and teacher, 1926-2019, born in Argentina) [500023533]

founder of .... [Eero Saarinen & Associates](#)

..... (American architectural firm, active 1950-1961) [500119694]

member of .... [Saarinen, Saarinen and Associates](#)

..... (American architectural partnership, active 1947-1950) [500229797]

member of .... [Saarinen, Swanson, Saarinen](#)

..... (American architectural partnership, active 1941-1947) [500229808]

partner of .... [Saarinen, Eliel](#) 1941-1950

..... (Finnish architect, 1873-1950, active in the United States) [500027014]

partner of .... [Swanson, J. Robert F.](#) 1941-1947

..... (American architect, active late 20th century) [500113110]

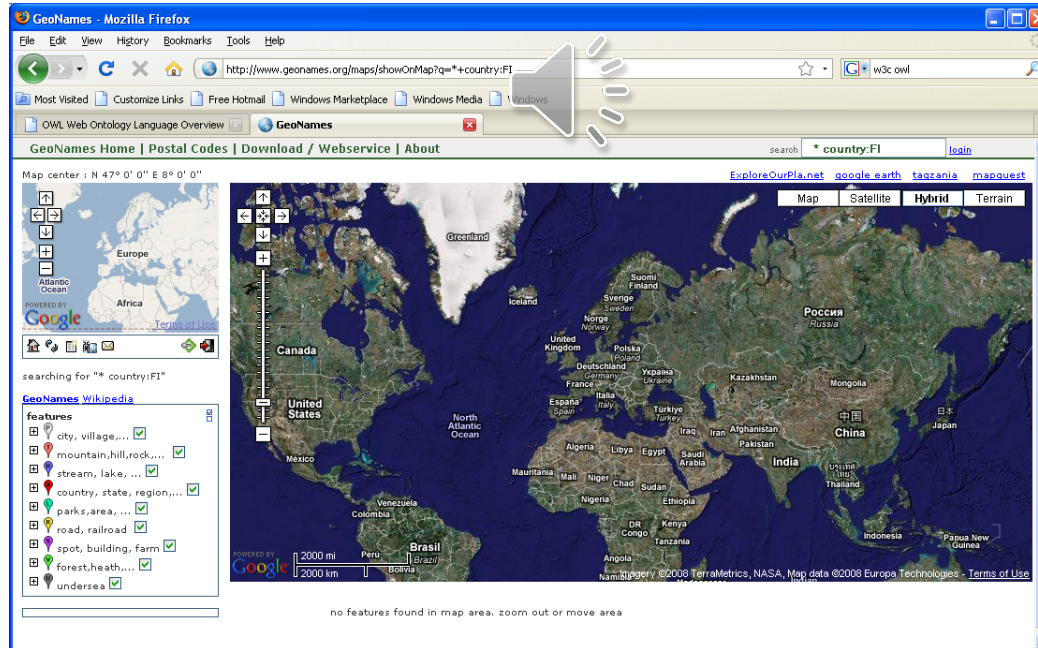
related to .... [Saarinen and Saarinen](#)

..... (Finnish architectural firm, contemporary) [500291347]



# Geonames

- Classes: 9 feature classes, 645 feature codes
- Instances:
  - *8 million geographical names, 6.5 million unique features, 2.2 million populated places, 1.8 million alternate names*
  - *Registries and Wiki used for populating the ontology*



# TGN Thesaurus of Geographical Names

- 912,000 records
- 1.1 million names, place types, coordinates, and descriptive notes
- Places important for the study of art and architecture
- Available in a Linked Open Data service: [Getty Thesaurus of Geographic Names \(Getty Research Institute\)](#)



The screenshot shows a web browser window titled "TGN Full Record Display, English (Getty Research) - Mozilla Firefox". The address bar contains the URL: <http://www.getty.edu/vow/TGNFullDisplay/Find=helsinki&place=&nation=>. The page content includes the following sections:

- Research** header with navigation links: [Research Home](#) > [Conducting Research](#) > [Thesaurus of Geographic Names](#) > [Full Record Display](#)
- Getty Thesaurus of Geographic Names® Online Full Record Display**
- Buttons for [New Search](#), [Previous Page](#), and [Help](#).
- Language options: [Vernacular Display](#) | [English Display](#)
- Instruction: "Click the icon to view the hierarchy."
- ID: 7009880** **Record Type:** administrative
- Helsinki (inhabited place)**
- Coordinates:**  
Lat: 60 08 00 N *degrees minutes*    Lat: 60.1333 *decimal degrees*  
Long: 025 00 00 E *degrees minutes*    Long: 25.0000 *decimal degrees*
- Note:** Located on peninsula in Gulf of Finland; founded by King Gustav Vasa of Sweden in the 16th century to compete with Reval (Tallinn), Estonia across the gulf; ravaged by plague in 1710, and fires in 1710 and 1808; prospered after being strongly fortified in 1748.
- Names:**  
**Helsinki (preferred, C, V, N)**  
**Helsingfors (C, V, N)**
- Hierarchical Position:**
  - World (facet)
  - ... Europe (continent)
  - ..... Finland (nation)
  - ..... Etelä-Suomen Lääni (province)
  - ..... Helsinki (inhabited place)
- Place Types:**
  - inhabited place (preferred, C) ..... founded in 1550 on the Vantaa estuary, moved to present site S of the estuary in 1640
  - city (C)
  - national capital (C)
  - port (C)
  - provincial capital (H) ..... of the former province of Uusimaa until 1998
- Sources and Contributors:**
  - Helsingfors ..... [BHA, VP]
  - ..... Times Atlas of the World (1994) 83
  - Helsinki ..... [BHA, FDA, GRLPSC, VP]
  - ..... Canby, Historic Places (1984) II, 385
  - ..... Columbia Lippincott Gazetteer (1961)
  - ..... Encyclopaedia Britannica (1988) V, 820
  - ..... Times Atlas of the World (1994) 83
  - ..... USBGN Bulletin 16 (1998) 3
  - ..... Webster's Geographical Dictionary (1984)

# W3C Standards for Semantic Web Ontologies/Vocabularies

## RDF Schema

- Class and property hierarchies

## SKOS Simple Knowledge Organization System

- Light-weight semantics
- E.g., for representing existing glossaries, thesauri, and classifications

## OWL Web Ontology Language

- Rich semantics based on logic
- Supports more advanced reasoning

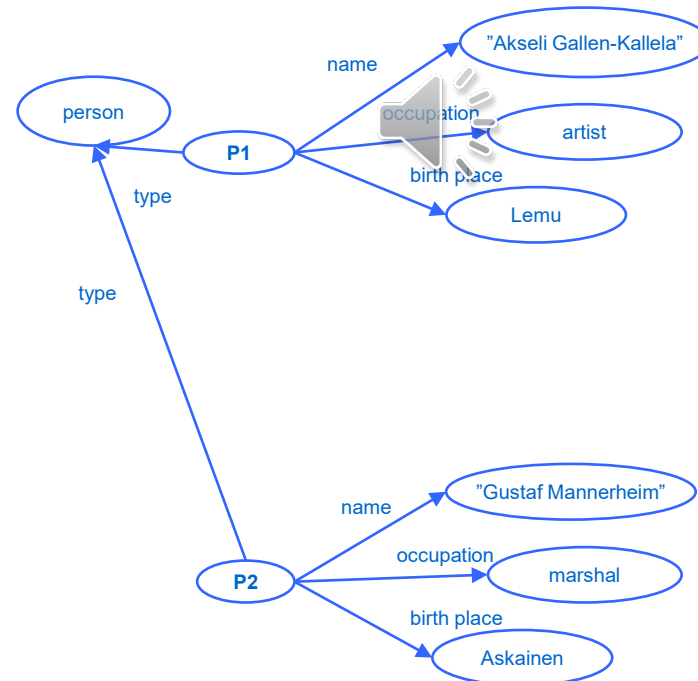
**Metadata + Ontologies =  
Linked Data (Web of Data)**



**Enriching Data by Data Linking  
through Shared Ontologies:  
An Example**

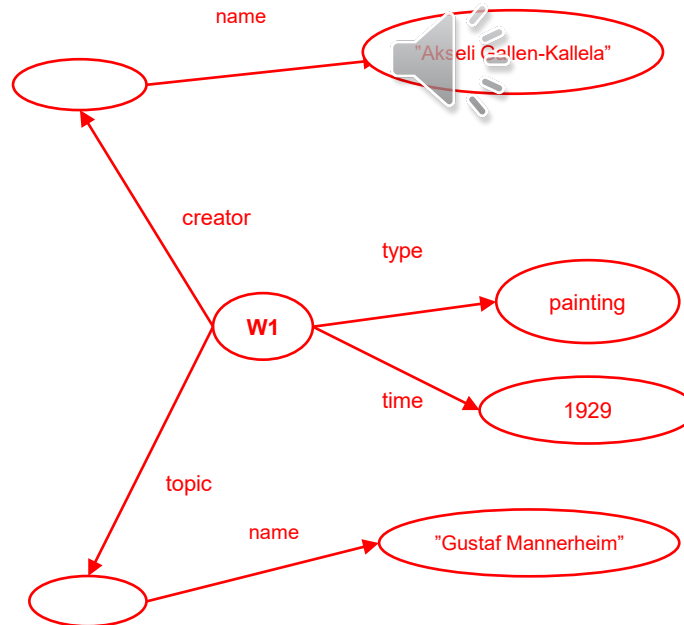
# Biography Centers and Libraries Have Databases of about Historical People

person	name	occupation	birth place	...
P1	Akseli Gallen-Kallela	artist	Lemu	
P2	Gustaf Mannerheim	marshal	Askainen	
...				



# Museums Catalogue Paintings

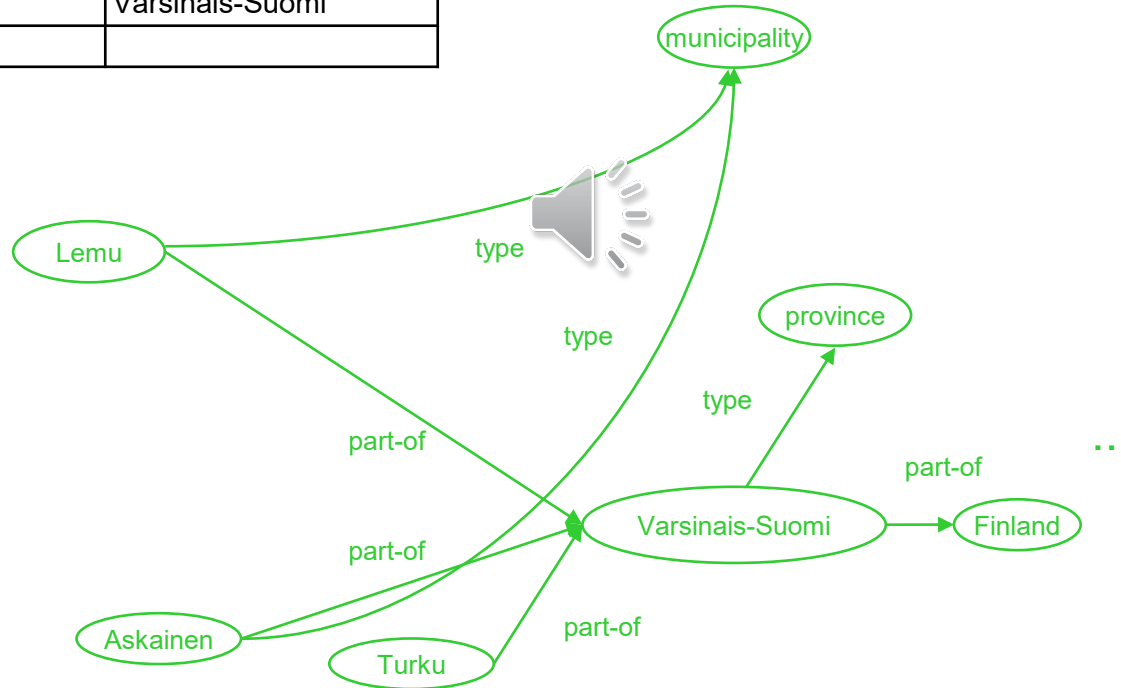
Work	name	creator	time	Topic	...
W1	Portrait of Mannerheim	Akseli Gallen-Kallela	1929	Gustaf Mannerheim	
W2	Aino Triptych	Akseli Gallen-Kallela	1891	Aino, Kalevala	
...					





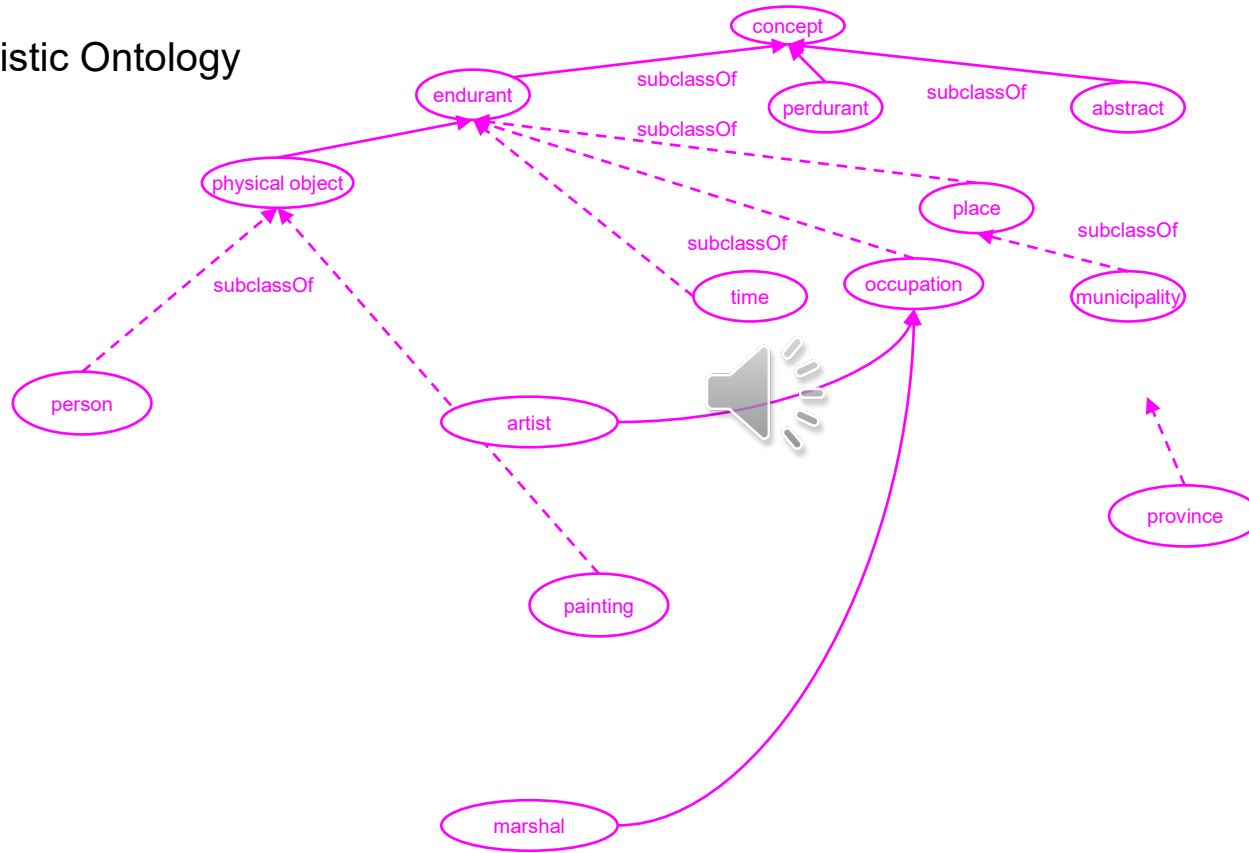
# Land Surveys Maintains Place Registries

municipality	province
Askainen	Varsinais-Suomi
Helsinki	Uusimaa
Lemu	Varsinais-Suomi
Turku	Varsinais-Suomi
...	

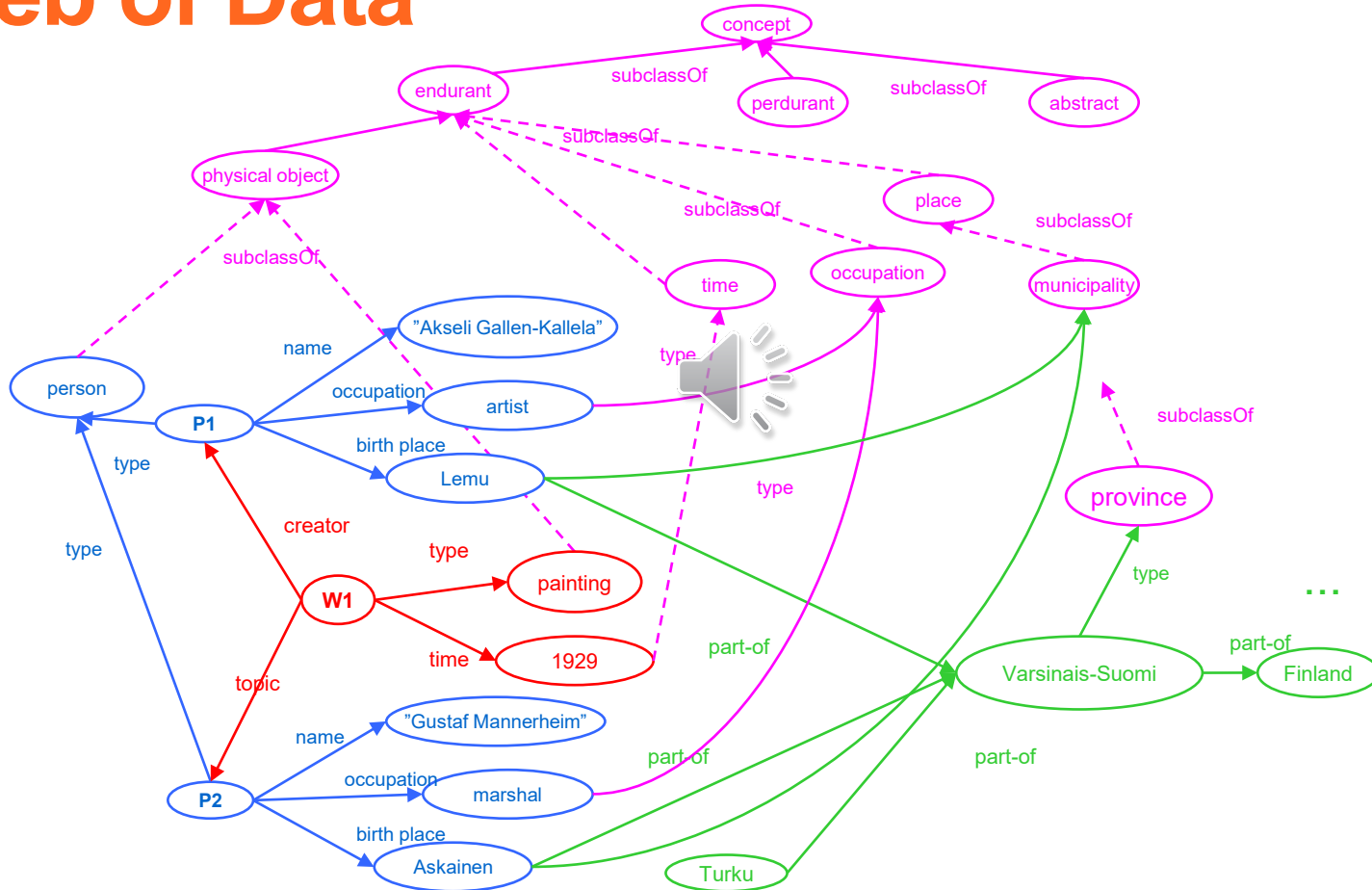


# National Library Builds Ontologies

Holistic Ontology



# Semantic RDF Graph Combines All Data: Web of Data



# Two Key Challenges in Aggregating Data by Data Linking

**Ontologies** used is metadata descriptions must be **shared** by collaborating parties

- Otherwise the data just does not link properly!
- Multiple concepts for the same thing emerge



**Metadata models** have to be **aligned**

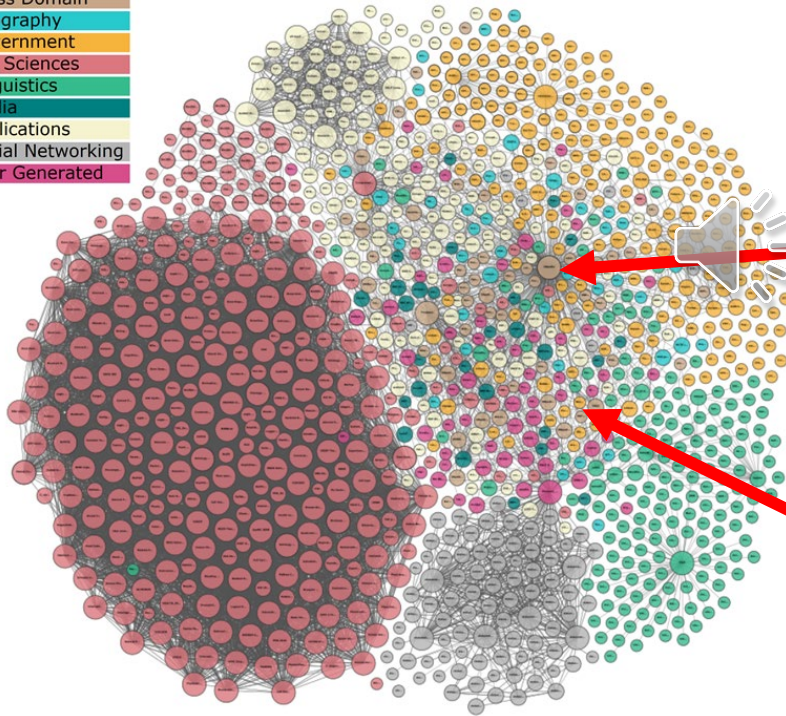
- E.g. two libraries providing data about books in different forms
- Otherwise the data is not interoperable

# Web of Data: Linked Open Data Cloud

## Human Knowledge on the Semantic Web

### Legend

Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated



**Wikipedia**

**WarSampo**

**LODStats 2018 (<http://stats.lod2.eu/>):  
10 000 datasets, 150 billion triples**

# Application Example: WarSampo – Finnish WW2 on the Semantic Web



- <https://vimeo.com/212249404>

# Rule Level - Logic



# The Idea of Rules

- Semantic web semantics is based on **logic**
- Logic = “new” information can be derived from old by (rule-based) reasoning





# Rule Markup Language RuleML

## Standardized XML notation for rules

$\text{hasParent}(?x1, ?x2) \wedge \text{hasBrother}(?x2, ?x3) \Rightarrow \text{hasUncle}(?x1, ?x3)$

```
<ruleml:imp>
  <ruleml:rlab ruleml:href="#example1"/>
  <ruleml:_body>
    <swrlx:individualPropertyAtom swrlx:property="hasParent">
      <ruleml:var>x1</ruleml:var>
      <ruleml:var>x2</ruleml:var>
    </swrlx:individualPropertyAtom>
    <swrlx:individualPropertyAtom swrlx:property="hasBrother">
      <ruleml:var>x2</ruleml:var>
      <ruleml:var>x3</ruleml:var>
    </swrlx:individualPropertyAtom>
  </ruleml:_body>
  <ruleml:_head>
    <swrlx:individualPropertyAtom swrlx:property="hasUncle">
      <ruleml:var>x1</ruleml:var>
      <ruleml:var>x3</ruleml:var>
    </swrlx:individualPropertyAtom>
  </ruleml:_head>
</ruleml:imp>
```



# Application Example: MuseumFinland Recommends

## Inference rules tell machine about the world

- E.g., that "student's cap" is related to "parties"
- E.g., that entities are related to each other if their superclasses are related to each other

## Based on the graph of metadata + ontologies, machine can:

- Reason interesting new relations between museum items, and
- Provide them to end users as recommendation links

# Application example: MuseumFinland



The screenshot shows the MuseumSuomi website in Internet Explorer. The page title is "MuseoSuomi - Suomen museot semanttisessa webissä". The search results show a single item: "Pullonsuojus, 2 kpl:istuva koira". The item details include:

- Materiaali:** viinapullo: lasi, pulonsuojus: lanka
- Valmistaja:** Karhulan lasitehdas, Tapio Wirkkala
- Valmistusaika:** 1962, 1970-1 n.
- Valmistustekniikka:** viinapullo: tehtäminen, pulonsuojus: käsityötä
- Käyttäjät:** Eero Kallio
- Käyttöpaikka:** Etelä-Suomen lääni, Suomi
- Asiasana:** ALKOHOLJUOMAT, ELÄINHAHMOT, KORISTE-ESINEET
- Mitat:** pullon pohjan halkaisija 6,5cm, korkeus 22,5cm, pulonsuojuksen korkeus 29,0cm
- Museokokoelma:** LAHDEN HISTORIAALINEN MUSEO
- Vastuumuseo:** LAHDEN KAUPUNGINMUSEO
- Asiasanasto:** Lahden kaupunginmuseon sanasto
- Eseen numero:** LKMLHMLHMES.95073:154
- ID:** 95073154
- Viinapullo:** Alkon Koskenkorvapullo. Lieriömäinen, loivat hartiat. Korot ja etiketti puuttuvat. Pulonsuojus: istuvan koiran muotoinen pulonsuojus. Muodossa kahdesta osasta: koiran vartalosta ja päästä. Koiran vartaloon on ommeltu viisi leikkatupsua (jalat ja häntä), ylhäällä lankakiristys. Koiran pää on virkattu talouspaperirullasta leikatun leriön ympärille. Kasvoissa mustat napit silminä, erillinen pieni kuono ja kolme lankatupsua (posket ja päällella oleva otsanukka).

The right sidebar contains several sections:

- Sama käyttäjä**
  - Eero Kallio:
    - [Keräilykortti, 14 kpl:tuotemainoskortti, erilaisia](#)
    - [Kulho, 4 kpl:jalkiruokakulho](#)
    - [Päähine, miehen: turkislakki, 'suikka'](#)
    - [Taskuliina, miehen: taskuliinan korvike](#)
    - [Jalkineet, miehen: koripallokengät](#)
- Samaan aiheeseen liittyviä esineitä**
  - alkoholijuoma:**
    - [kanisteri: taskumatti](#)
    - [kanisteri: taskumatti](#)
    - [kanisteri: taskumatti](#)
    - [viinipullo: lasipullo](#)
    - [pullo: lasipullo](#)
  - eläimet:**
    - [kuvakirja: kuvakirja, kangasta](#)
    - [helisin: purulelu](#)
    - [muovikarhu: vinkuva karhulelu](#)
    - [säästölipas: vanerilipas](#)
    - [malja: purvati](#)
  - koriste-esineet:**
    - [luonontausta: mietelausetauhu](#)
    - [lyhty: öljylamppu](#)
    - [kannu: koristekannu](#)
    - [kynntilänjalja](#)
    - [maljakko: maljakko](#)
- Sama materiaali**
  - lanka:**
    - [neuletakki: naisen neuletakki](#)
    - [tauhu: kehystetty kirjontavo](#)

# Application Domains of Semantic web

- **Semantic portals**
- **Information retrieval systems**
- **Recommender systems**
- **Knowledge management systems**
- **Personalized systems**



...

**Examples of applications / domains have been collected here:**

**<https://www.w3.org/2001/sw/sweo/public/UseCases/>**

# What is New? Components of Semantic Web

PROGRAMMING

Object-oriented  
modelling / ontologies

ARTIFICIAL  
INTELLIGENCE

Description logic  
semantics



XML syntax,  
RDF data model, ...

WWW TECHNOLOGIES

# What is the Semantic web?

**Content perspective:** A new metadata layer on the Web describing its contents in terms of shared vocabularies, i.e., ontologies

- Web as a global database system
- Web of Pages vs. Web of Data

**Application perspective:** Machine-understandable web

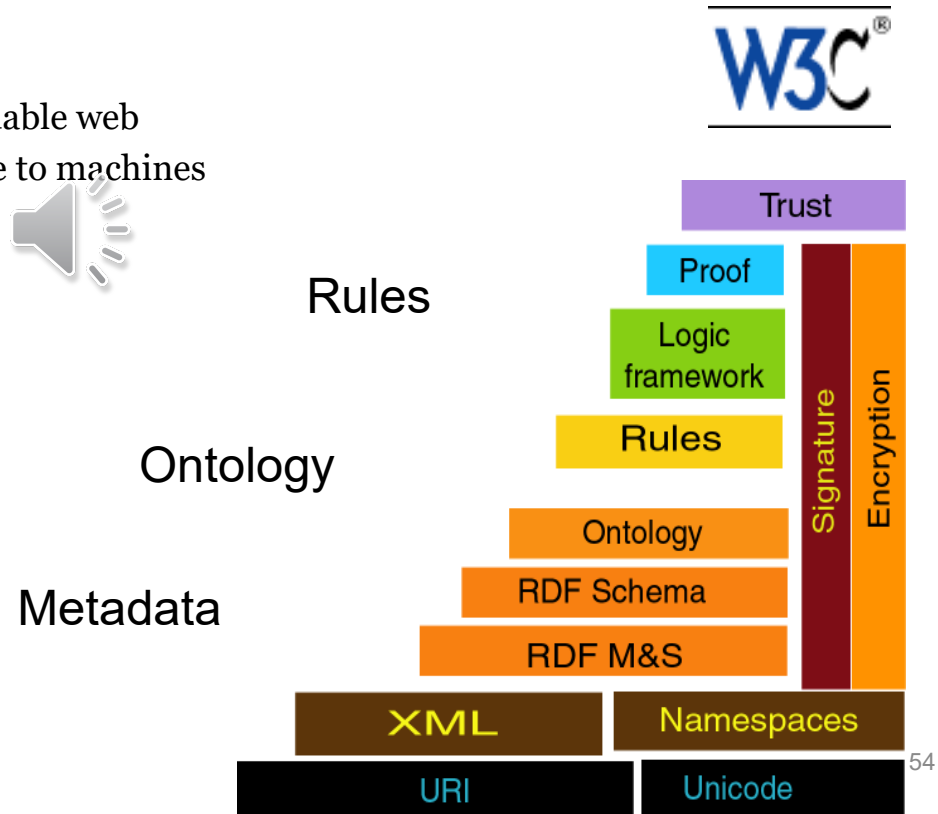
- The meaning (semantics) of contents accessible to machines
- Enables human usage
  - *Intelligent web services*
  - *Semantic interoperability*



**Technological perspective:**

Next layers above XML

- W3C standards:  
RDF(S), OWL , SPARQL, etc.



# Questions

