HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

Elements of a National Semantic Web Infrastructure

Case Study Finland on the Semantic Web

FINNONTO

Prof. Eero Hyvönen
Semantic Computing Research Group (SeCo)
Helsinki Univ. of Technology (TKK), Media Technology
University of Helsinki, Dept. of Computer Science
http://www.seco.hut.fi/

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Outline of Talk

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

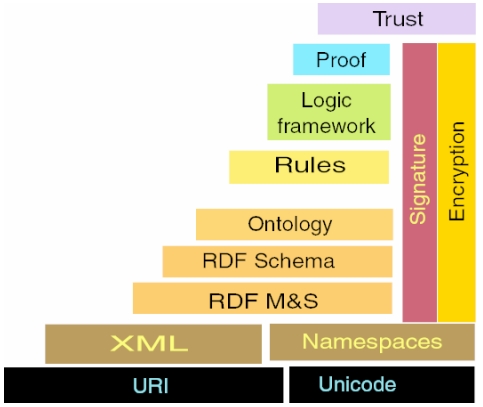
- The semantic web is coming
- A content infrastructure on the web is needed for it
- What should be done about it?
  - The vision of a national FinnONTO-project 2003-2007
  - Goal and results
    - » Standards
    - » Ontologies
    - » Ontology services
    - » Tools
    - » Applications: semantic web at work

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Perspectives of 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
- **Application perspective:** Machine understandable web
  - » The meaning (semantics) of contents accessible to machines
  - » Enables
    - Intelligent web services
    - Semantic interoperability
- **Techological perspective:** Next technological layers above XML
  - » W3C standards: RDF, OWL etc.



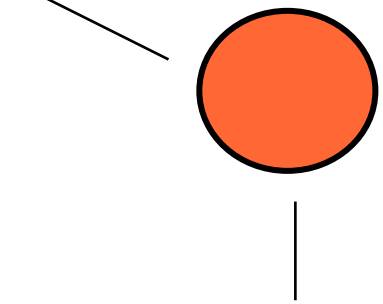
WHAT IS NEW?

PROGRAMMING

ARTIFICIAL INTELLIGENCE

Object-oriented modeling

Knowledge representation
Logic based semantics
Ontologies



XML, RDF(S), OWL,…

WWW-TECHNOLOGIES

A Climpse of RDF(S)

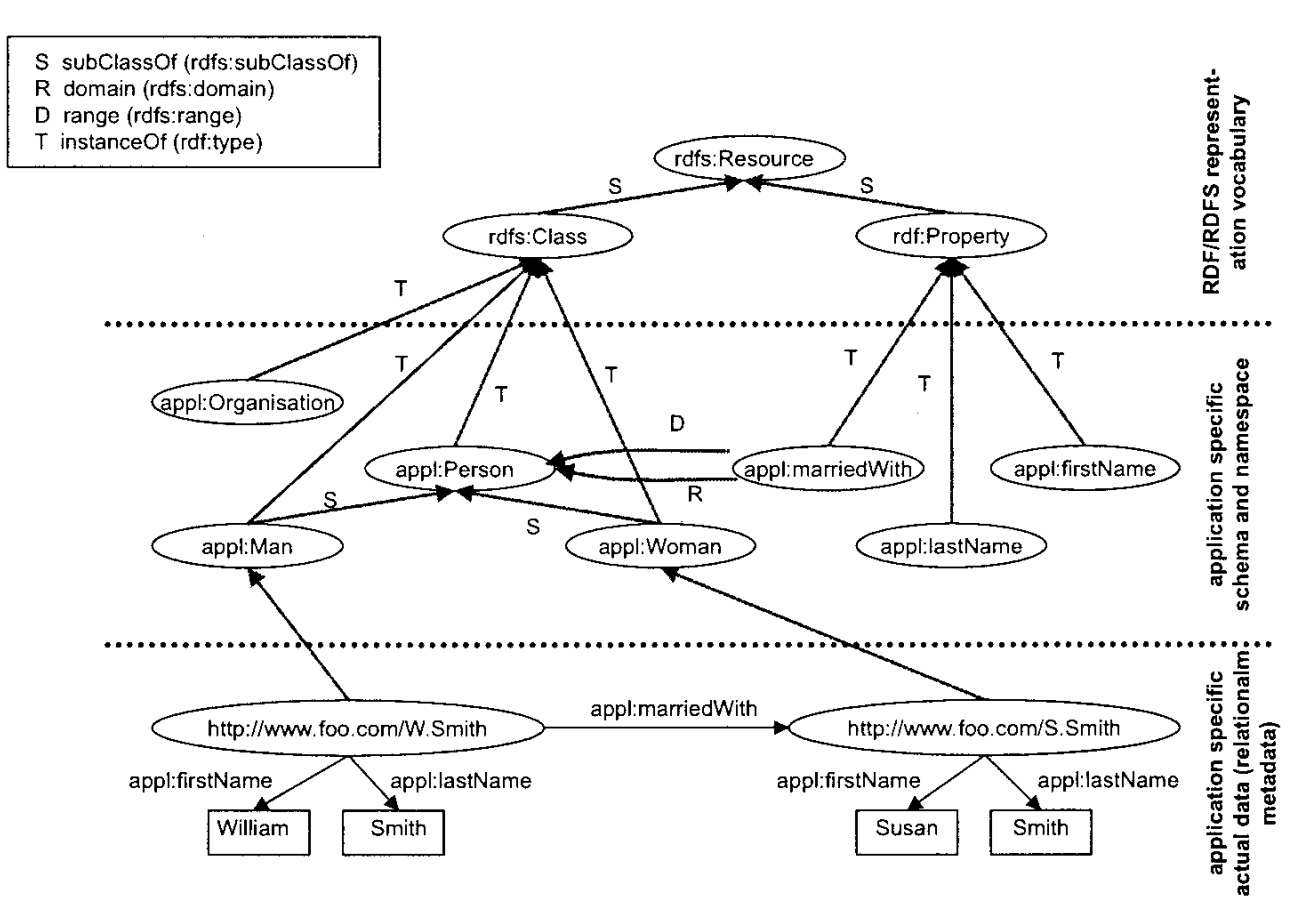


Figure 3.4. An RDF-Schema Example

(Maedche, 2002)

A National Problem

- Semantic Web = next generation/layer of the Web
- Ontologies = ”silver bullet” of the Semantic Web
- Finnish ontologies did not exist
- Something should be done about it!

- A solution approach:
  The National Finnish Semantic Web
  Ontology Project FinnONTO (2003-2007)

FinnONTO Thesis

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

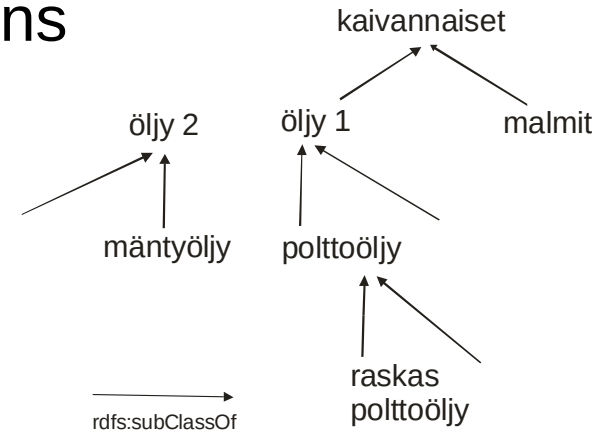
- Semantic Web needs a **content infrastructure**
  - Like traffic needs roads
  - Like energy service needs powerlines, power plants, standards, ...
  - Like mobile phones need GSM or 3G-networks

UNIVERSITY OF HELSINKI

SeCo
semantic computing

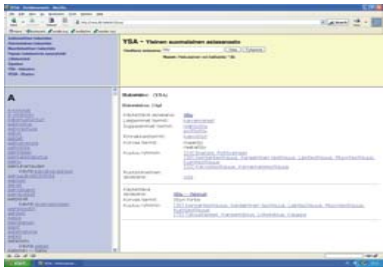
FinnONTO Solution Approach

- Major infrastructure components
  - **Ontologies** to be shared
  - **Ontology services** for utilizing ontologies
  - **Standards** to make things interoperable
    - » **E.g. metadata standards**
  - **Tools** to help in creating applications



The Way to Go

- Start a national multi-domain ontologization process
  - Making contents of different domain interoperable
  - Thesauri -> ontologies
    - » Human usage -> human/machine usage
  - Key ontologies should be open source and maintained publicly
    - » Wide acceptance and usage
- Business applications can be built effectively upon a solid infrastructure



FinnONTO project organization

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Universities as research partners
  - Helsinki University of Technology (TKK)
    - » Lab. of Media Technology and Lab. of Cartography & Geoinformatics
  - University of Helsinki (UH)
    - » Department of Computer Science and Dept. of General Linguistics
  - University of Tampere
    - » Department of information sciences
- Directed and 80-90% of the research done by the Semantic Computing Research Group
  - TKK and UH

UNIVERSITY OF HELSINKI

SeCo
semantic computing

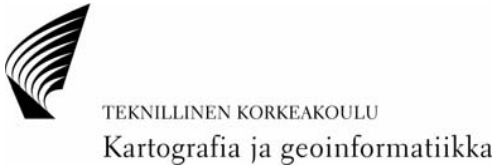
Industrial & public organization consortium

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- 2003-2004
  - 14 funding organizations
- 2004-2005
  - 16 funding organizations
- 2005-2006
  - 30 funding organizations
- 2006-2007:
  - 37 funding organizations
- Current budget 0,8M€ / year (Tekes 80%)

UNIVERSITY OF HELSINKI

SeCo
semantic computing



Antikvaria-ryhmä

Goals

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- 1. **Ontology development** open source
  - General Finnish Ontology based on the national YSA thesaurus (23,000 concepts)
  - Various vertical ontologies based on YSO
- 2. **ONKI ontology services**
  - Content indexing using ontology web services
  - Ontology-based information retrieval
  - Collaborative ontology development
- **3. Pilot applications & tools**
  - Eating our own dog food

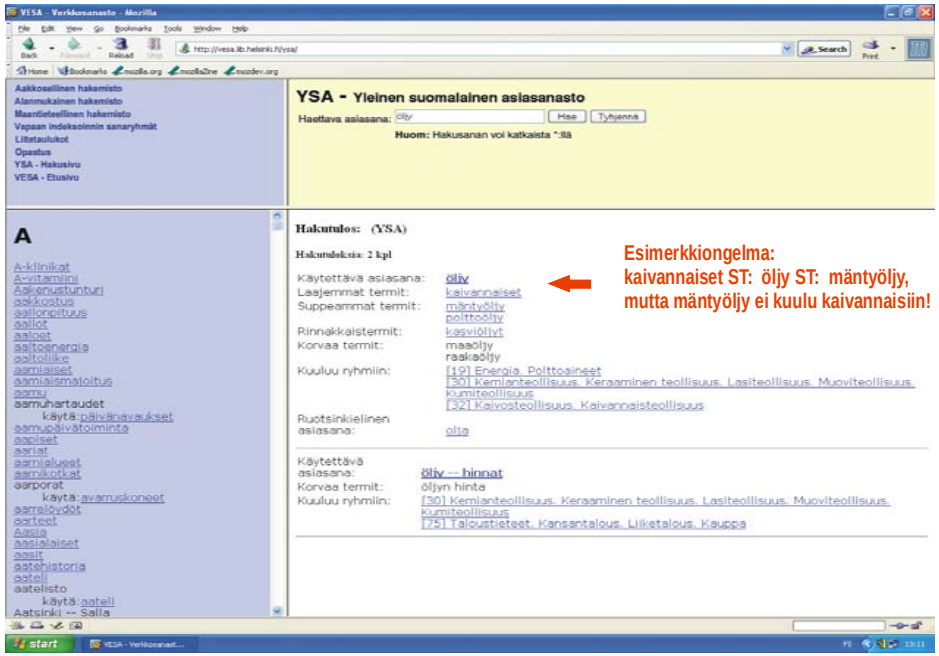
UNIVERSITY OF HELSINKI

SeCo
semantic computing

1. Ontology development

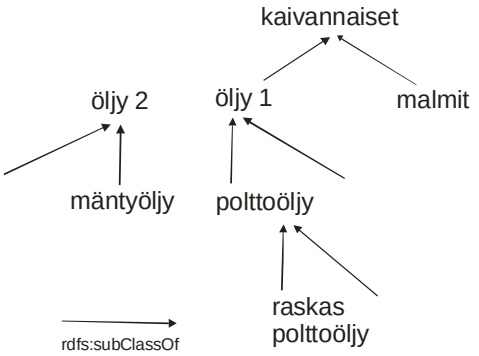
- Motto: Thesauri -> Ontologies!

YSA



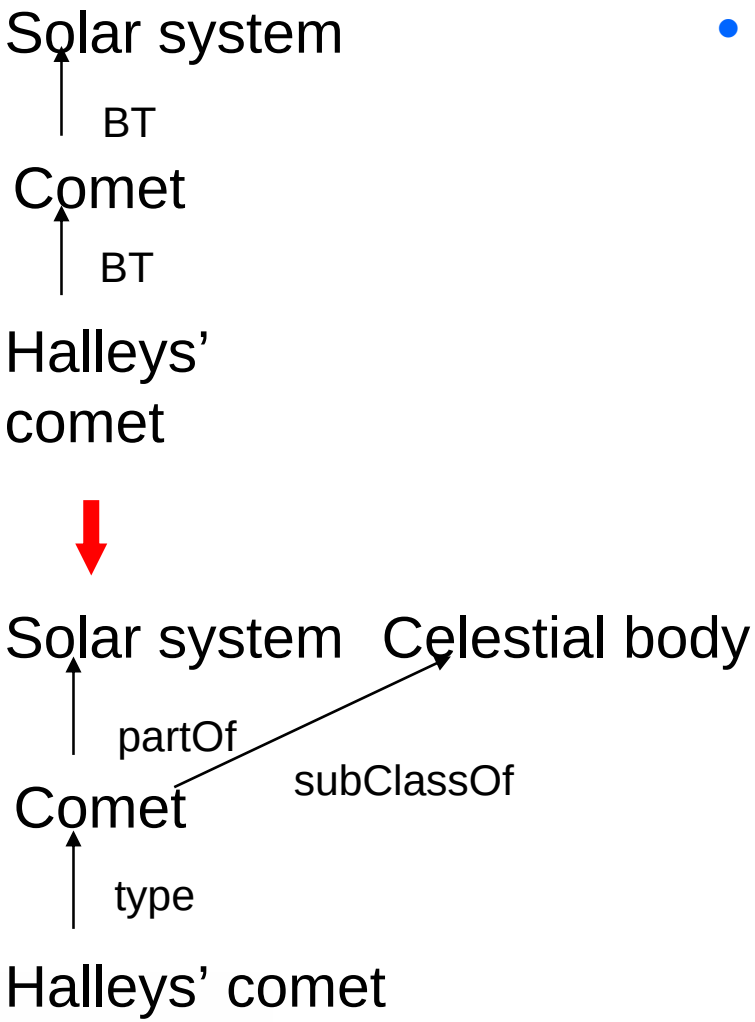
Esimerkkiongelma:
kaivannaiset ST: öljy ST: mäntyöljy,
mutta mäntyöljy ei kuulu kaivannaisiin!

YSO



Why Thesauri are Not Enough but Ontologies are needed?

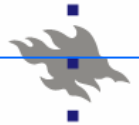
- Example from the YSA-thesaurus:



- The machine is confused:
  - Is Halley's comet an individual or a class of them, such as Comet?
  - Can there be many Halley's comets or only one?
  - Is Comet a kind of Solar system or a part of a solar system. Is it a part as a concept or are all individual comets a part of some solar system?
  - What does "part of" mean: real part of, contained in, member of, made of, connected to.
  - Do comets have properties of solar systems (e.g. own planets) based on BT
  - Searching "Solars systems" would retrieve comets although comets are not solar systems
  - ...

A Key Point of Ontologies: Using URIs, not Keywords!

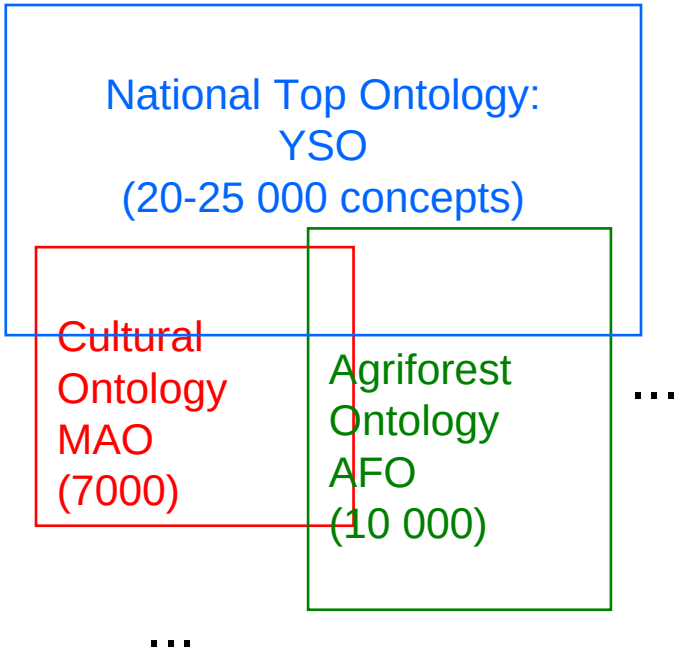
- Each concept will have a globally unique URI (across all domains)
  - URI = Universal Resource Identifier
    - » URL web addresses are a special case of URIs
  - A keyword is not enough for indexing the meaning:
    - » E.g. ”Nokia”:
      - = ”Nokia” as a company?
      - = ”Nokia” as a city in Finland?
      - = ”Nokia” a character in a F.E. Sillanpää’s novel?
      - = ”Nokia” as an animal?
    - » E.g. ”Pyhäjärvi” as a location
      - There 49 Pyhäjärvi lakes, villages etc. in Finland
- The URIs are globally shared among users



Idea: Horizontal top ontology + vertical domain ontologies

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Top ontology YSO as semantic glue
- Merges overlapping domain ontologies
  - Cultural ontology MAO
  - Place ontology SUO
  - Time-place ontology SAPO
  - Actor ontology TOIMO
  - Event & Process ontology TAO
  - Photography ontology VALO
  - Agriforest ontology AFO
  - Ontology of Finnish History HISTO
  - ...
- Alingning other classification systems and top ontologies
  - » HKLJ + YSO (library domain)
  - » ICONCLASS + YSO (fine arts domain)
  - » MeSH + YSO (medical domain)



UNIVERSITY OF HELSINKI

SeCo
semantic computing

Metadata standards

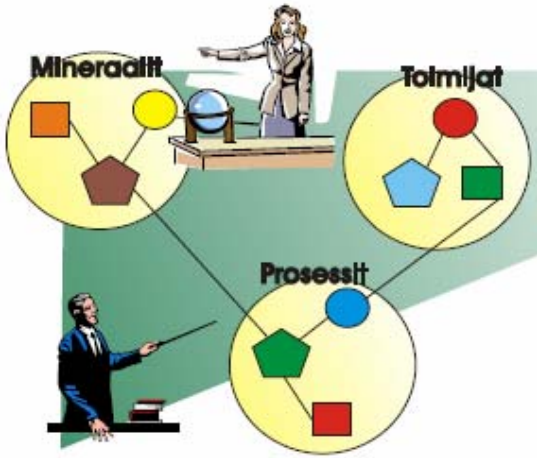
- Ontology structures
- Cultural content types
- Health promotion information and services
- Spatial metadata
  - With a national working group
- Learning objects
  - With FinnMeta group
- ...

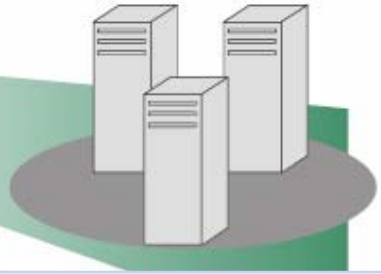
2. Ontology Services & User Groups

1. Ontology Developers
- Colloborative development of interdependent ontologies
- Versioning and support for updates

2. Information Searchers
- Support concept-based search
- Keyword disambiguation
- Finding the right search concepts

Nokia:
company or city?





2. Information Indexers
- Support indexing concept finding
- Keyword disambiguation
- Support indexing patterns

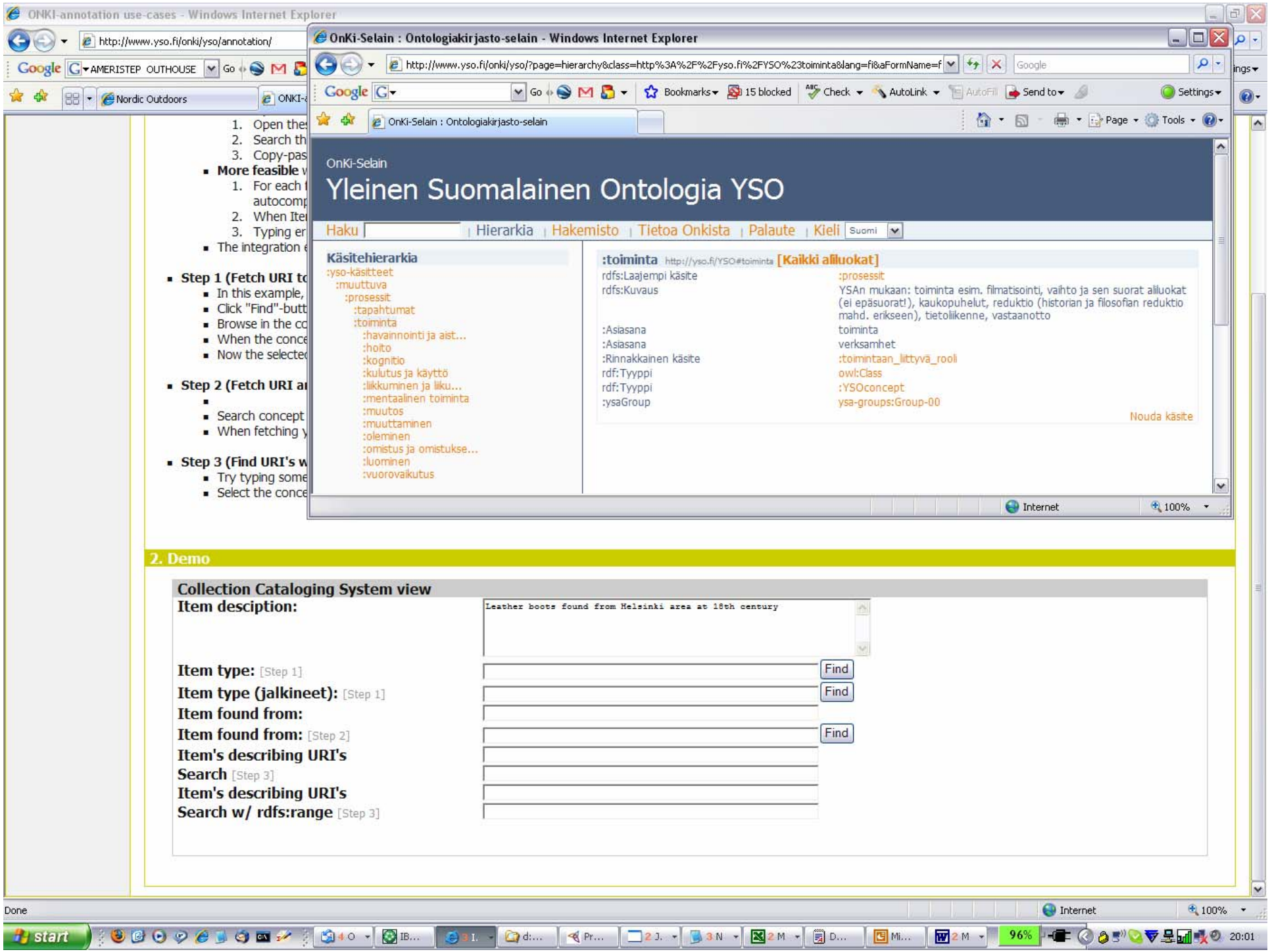
ONKI Ontology Service Demonstration

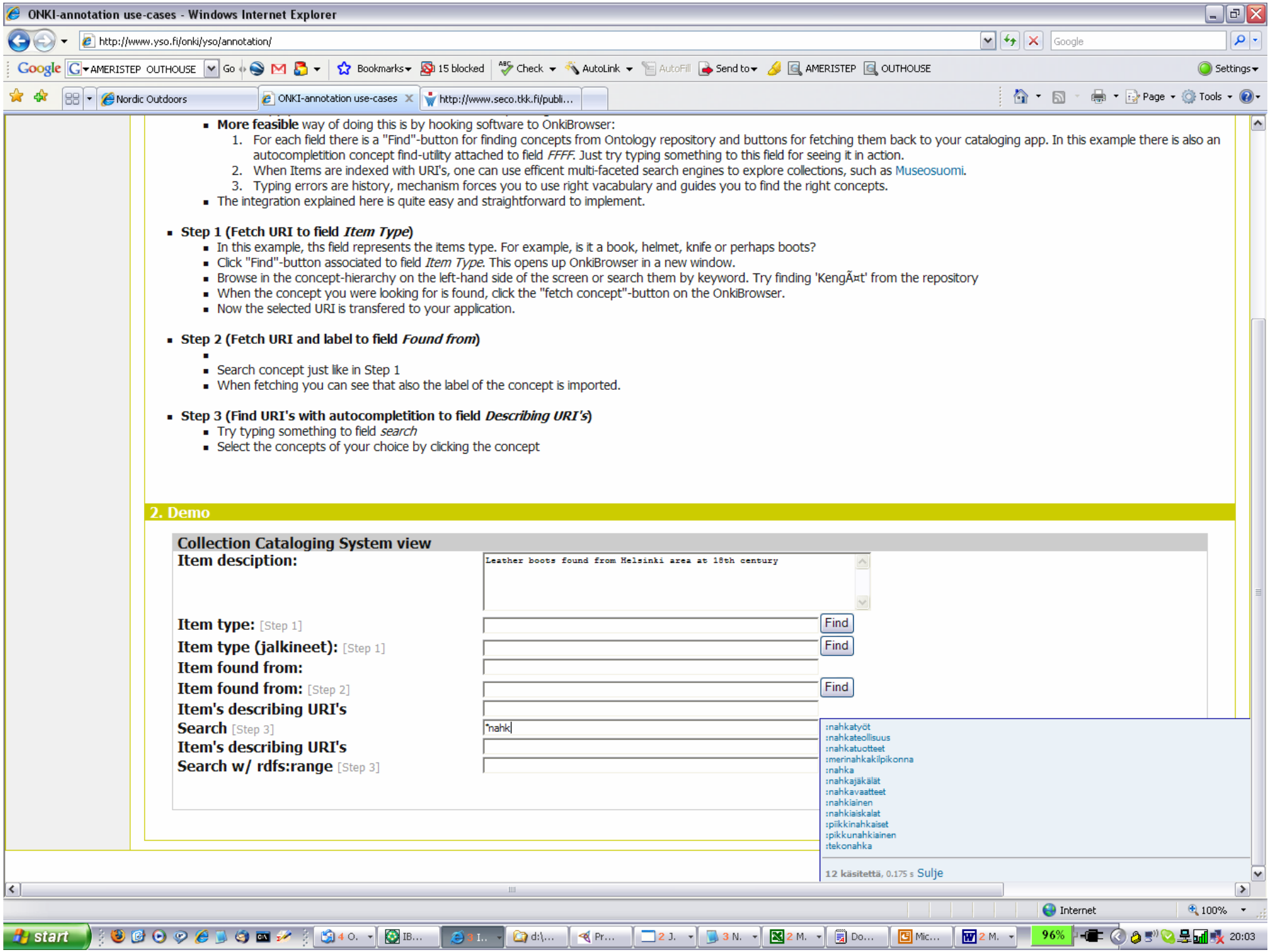
HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

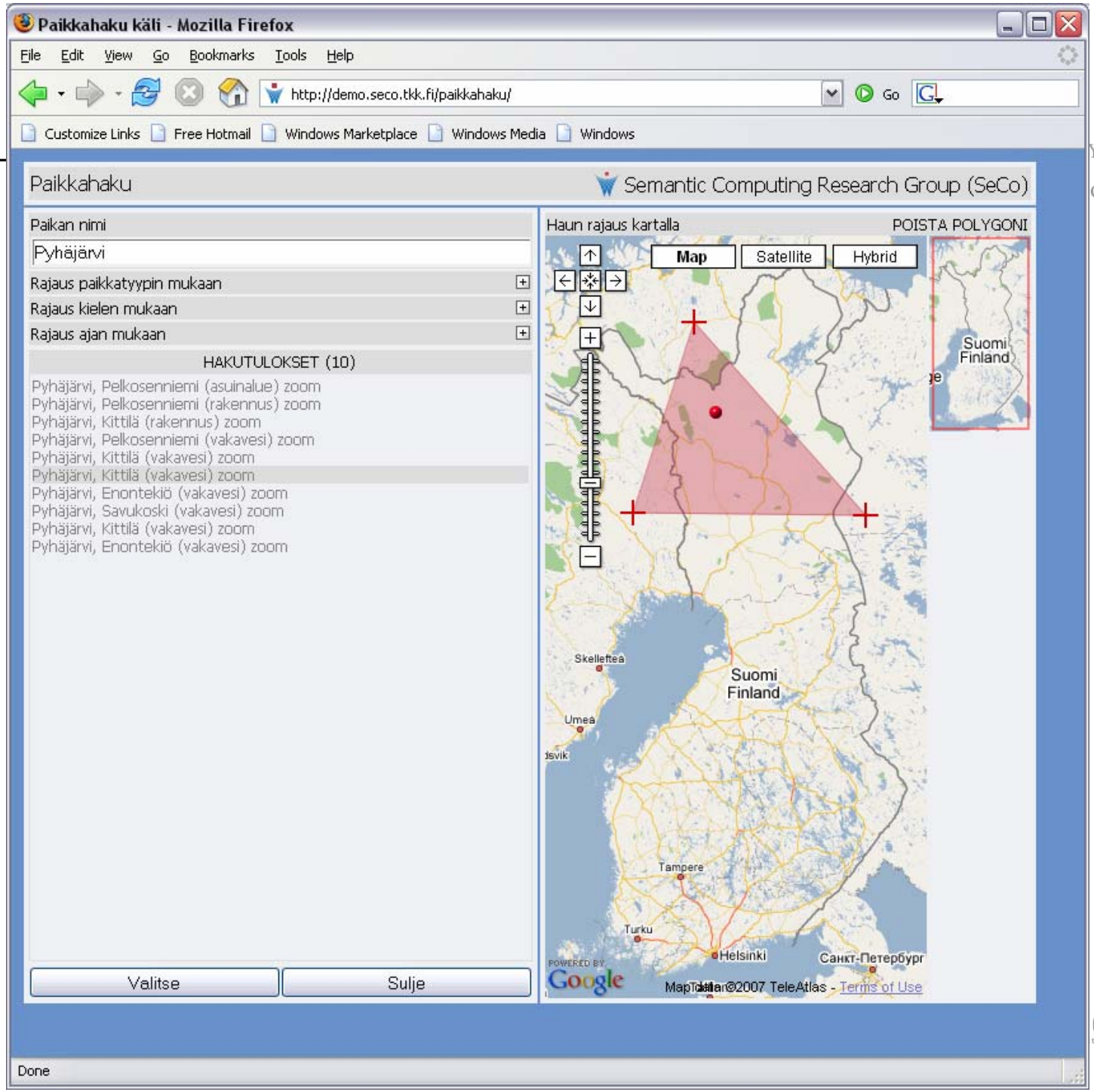
- Sharing ontologies on the web when indexing content
- Demonstration
  - Indexing with concepts (meaning), not with keywords
  - Finding the right indexing annotation concept
  - Retrieving the corresponding URI automatically to an external application
  - (Komulainen, Hyvönen, Valo, DC 2005)

UNIVERSITY OF HELSINKI

SeCo
semantic computing


ONKI-annotation use-cases - Windows Internet Explorer
http://www.yso.fi/onki/yso/annotation/
OnKi-Selain : Ontologiakirjasto-selain - Windows Internet Explorer
http://www.yso.fi/onki/yso/?page=hierarchy&class=http%3A%2F%2Fyso.fi%2FYSO%23toiminta&lang=fi&aFormName=f
OnKi-Selain : Ontologiakirjasto-selain
OnKi-Selain
Yleinen Suomalainen Ontologia YSO
Haku
Hierarkia
Hakemisto
Tietoa Onkista
Palaute
Kieli
Suomi
Käsitehierarkia
:yso-käsitteet
:muuttuva
:prosessit
:tapahtumat
:toiminta
:havainnointi ja aist...
:hoito
:kognitio
:kulutus ja käyttö
:liikkuminen ja liiku...
:mentaalinen toiminta
:muutos
:muuttaminen
:oleminen
:omistus ja omistukse...
:luominen
:vuorovaikutus
:toiminta http://yso.fi/YSO#toiminta [Kaikki aliluokat]
rdfs:Laajempi käsite :prosessit
rdfs:Kuvaus YSAn mukaan: toiminta esim. filmatisointi, vaihto ja sen suorat aliluokat (ei epäsuorat!), kaukopuhelut, reduktio (historian ja filosofian reduktio mahd. erikseen), tietoliikenne, vastaanotto
:Asiasana toiminta
:Asiasana verksamhet
:Rinnakkainen käsite :toimintaan_liittyvä_rooli
rdf:Tyyppi owl:Class
rdf:Tyyppi :YSOconcept
:ysaGroup ysa-groups:Group-00
Nouda käsite
Internet
100%
2. Demo
Collection Cataloging System view
Item desciption:
Leather boots found from Helsinki area at 18th century
Item type: [Step 1]
Item type (jalkineet): [Step 1]
Item found from:
Item found from: [Step 2]
Item's describing URI's
Search [Step 3]
Item's describing URI's
Search w/ rdfs:range [Step 3]
Find
Done
start
20:01


ONKI-annotation use-cases - Windows Internet Explorer
http://www.yso.fi/onki/yso/annotation/
Nordic Outdoors
ONKI-annotation use-cases
http://www.seco.tkk.fi/publi...
More feasible way of doing this is by hooking software to OnkiBrowser:
1. For each field there is a "Find"-button for finding concepts from Ontology repository and buttons for fetching them back to your cataloging app. In this example there is also an autocompletition concept find-utility attached to field FFFF. Just try typing something to this field for seeing it in action.
2. When Items are indexed with URI's, one can use efficent multi-faceted search engines to explore collections, such as Museosuomi.
3. Typing errors are history, mechanism forces you to use right vacabulary and guides you to find the right concepts.
The integration explained here is quite easy and straightforward to implement.
Step 1 (Fetch URI to field Item Type)
In this example, ths field represents the items type. For example, is it a book, helmet, knife or perhaps boots?
Click "Find"-button associated to field Item Type. This opens up OnkiBrowser in a new window.
Browse in the concept-hierarchy on the left-hand side of the screen or search them by keyword. Try finding 'KengÃ¤t' from the repository
When the concept you were looking for is found, click the "fetch concept"-button on the OnkiBrowser.
Now the selected URI is transfered to your application.
Step 2 (Fetch URI and label to field Found from)
Search concept just like in Step 1
When fetching you can see that also the label of the concept is imported.
Step 3 (Find URI's with autocompletition to field Describing URI's)
Try typing something to field search
Select the concepts of your choice by clicking the concept
2. Demo
Collection Cataloging System view
Item desciption:
Leather boots found from Helsinki area at 18th century
Item type: [Step 1]
Find
Item type (jalkineet): [Step 1]
Find
Item found from:
Item found from: [Step 2]
Find
Item's describing URI's
Search [Step 3]
*nahk
Item's describing URI's
Search w/ rdfs:range [Step 3]
:nahkatyöt
:nahkateollisuus
:nahkatuotteet
:merinahkakilpikonna
:nahka
:nahkajäkälät
:nahkavaatteet
:nahkiainen
:nahkiaiskalat
:piikkinahkaiset
:pikkunahkiainen
:tekonahka
12 käsitettä, 0.175 s Sulje
Internet
100%
start
20:03


Paikkahaku käli - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://demo.seco.tkk.fi/paikkahaku/
Go
Customize Links Free Hotmail Windows Marketplace Windows Media Windows
Paikkahaku
Semantic Computing Research Group (SeCo)
Paikan nimi
Pyhäjärvi
Rajaus paikkatyypin mukaan
Rajaus kielen mukaan
Rajaus ajan mukaan
HAKUTULOKSET (10)
Pyhäjärvi, Pelkosenniemi (asuinalue) zoom
Pyhäjärvi, Pelkosenniemi (rakennus) zoom
Pyhäjärvi, Kittilä (rakennus) zoom
Pyhäjärvi, Pelkosenniemi (vakavesi) zoom
Pyhäjärvi, Kittilä (vakavesi) zoom
Pyhäjärvi, Kittilä (vakavesi) zoom
Pyhäjärvi, Enontekiö (vakavesi) zoom
Pyhäjärvi, Savukoski (vakavesi) zoom
Pyhäjärvi, Kittilä (vakavesi) zoom
Pyhäjärvi, Enontekiö (vakavesi) zoom
Valitse
Sulje
Haun rajaus kartalla
POISTA POLYGONI
Map
Satellite
Hybrid
Suomi
Finland
Skellefteå
Umeå
Tampere
Turku
Helsinki
Санкт-Петербург
Google
©2007 TeleAtlas - Terms of Use
Done

Try ONKI Yourself

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Versions of new ONKI with
  - Various multi-lingual ontologies
  - Connected to ”legacy cataloging system” demo
  - Two lines of JavaScript is needed for an Ajax mash-up!
  - http://www.seco.tkk.fi/services/onki/

UNIVERSITY OF HELSINKI

SeCo
semantic computing

3. Pilot Applications

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- eCulture
  - MuseumFinland – Finnish Museums on the Semantic Web (2004)
  - CultureSampo – Finnish Culture on the Semantic Web (2005, 2006, 2007)
- eHealth
  - Semantic prototype of national citizens' health promotion portal Tervesuomi.fi (2007)
- eLearning
  - Orava – Semantic video & learning object portal (2004)
  - Opintie – General learning object portal (2007)
- eGovernment
  - Semantic citizen's Suomi.fi portal (2005)
- Tools
- Other topics

UNIVERSITY OF HELSINKI

SeCo
semantic computing

”MuseumFinland—Finnish Museums on the Semantic Web”

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Ideas
  - Global seamless view to heterogeneous collections
  - Semantic search + browsing
  - Common publication channel for museums
- (Inter)nationally awarded application
  - Semantic Web Challenge Award 2004 (2. prize)
  - Nordic Digital Excellence in Museums 2004 (nominee)
  - Prime Minister’s Innovation Acknowledgement

http://www.museosuomi.fi

Google Maps + MuseumFinland

(Hyvönen et al., JWS 2005)

UNIVERSITY OF HELSINKI

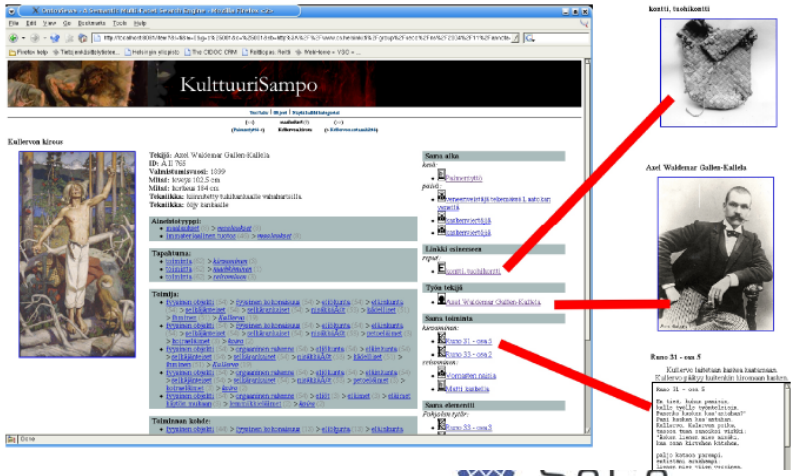
SeCo
semantic computing

”CultureSampo – Finnish Culture on the Semantic web”

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Generalizing MuseumFinland to cross-domain cultural content
- Modeling cultural context, events, and processes for enhanced intelligence
- Deep event-based semantic annotations
- New interface ideas: semantic autocompetion, maps, time lines, visualizing semantic structures
- New search ideas: searching and explaining links, meta-search
- Successive prototypes
  - 2005, 2006, 2007

(Junnila et al., FAIS 2006)
(Hyvönen et al., FAIS 2006)
(Kauppinen et al., FAIS 2006)



UNIVERSITY OF HELSINKI

SeCo
semantic computing

TerveSuomi.fi (”HealthyFinland.fi”)
The citizen’s semantic health promotion portal

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

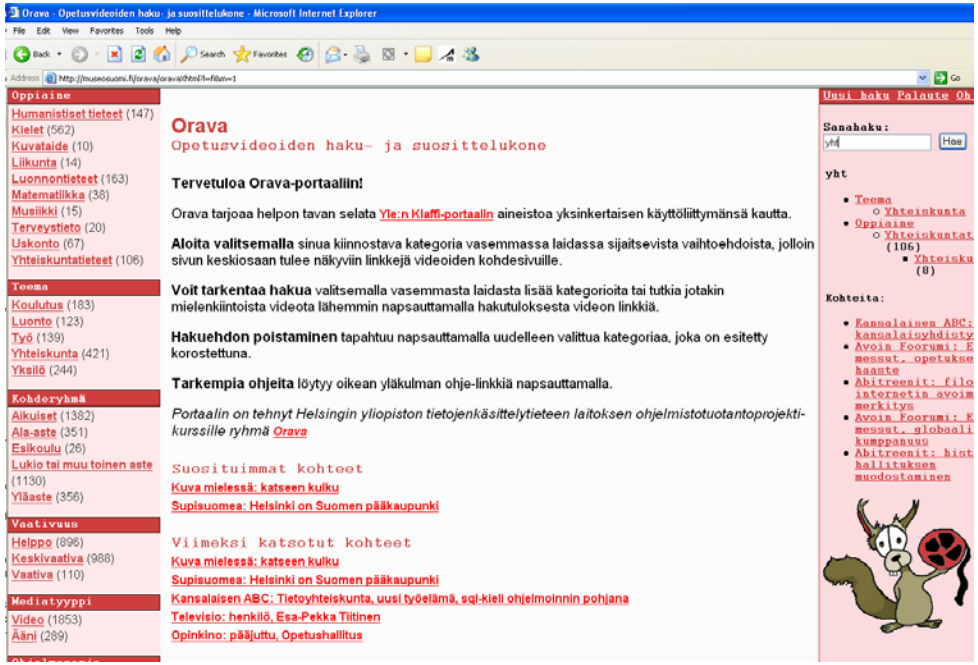
- Aggregates content produced by different Finnish health organizations
- Provides information through user-centric views
- Recommends related material of interest

- (Holi et al., ASWC 2006)
  (Suominen et al., ESWC 2007)
  (Hyvönen et al., ISWC 2007)

Orava: Video clip & learning object portal

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Semantic search & browsing
  - » 2200 videos, Learning Object Metadata (LOM)
- Semantic autocompletion
- Inter-portal semantic linking
  - » Linked with MuseumFinland



http://www.museosuomi.fi/orava

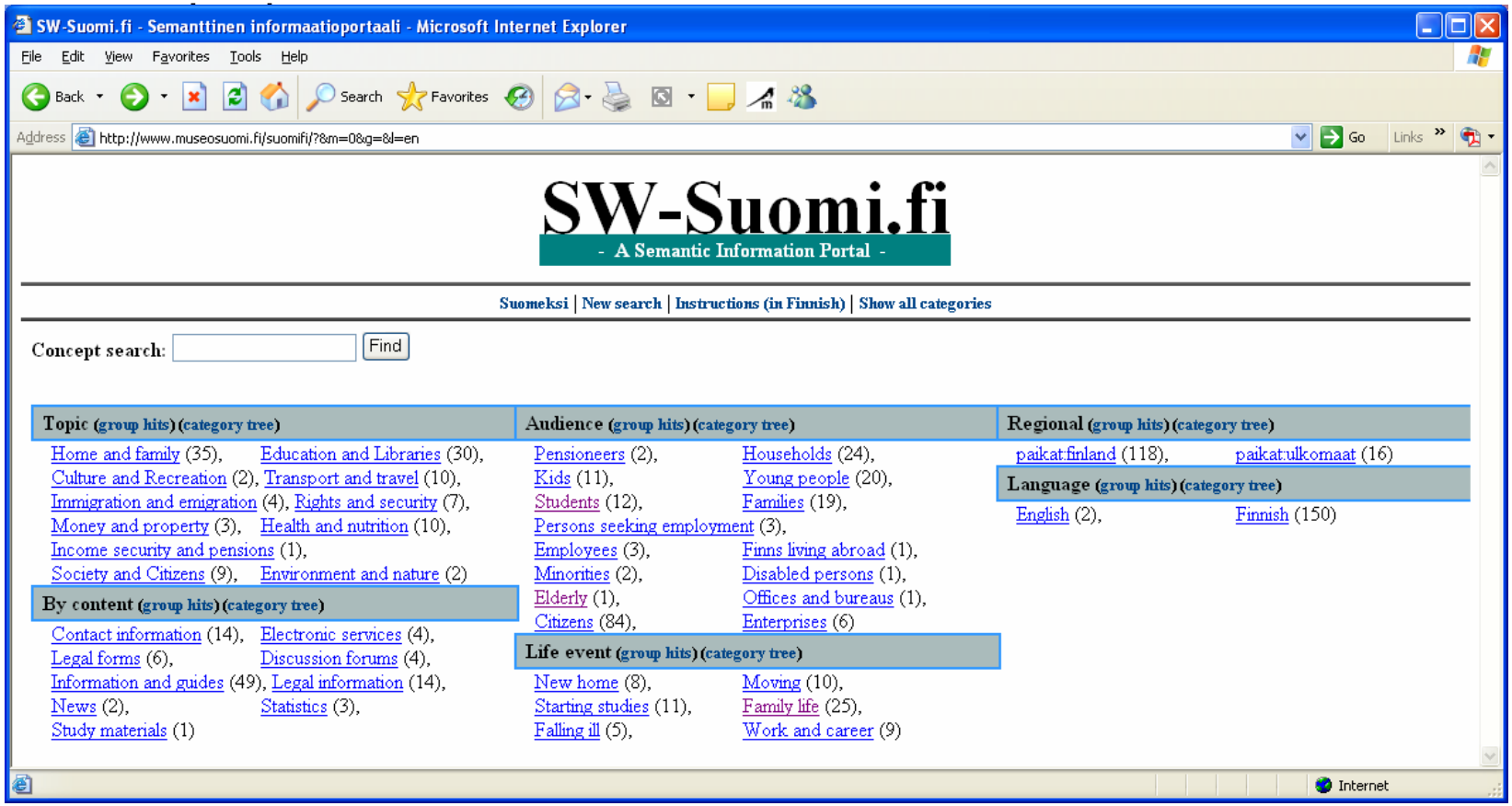
(Känsälä & Hyvönen, ASWC 2006 WS)

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Semantic Suomi.fi portal

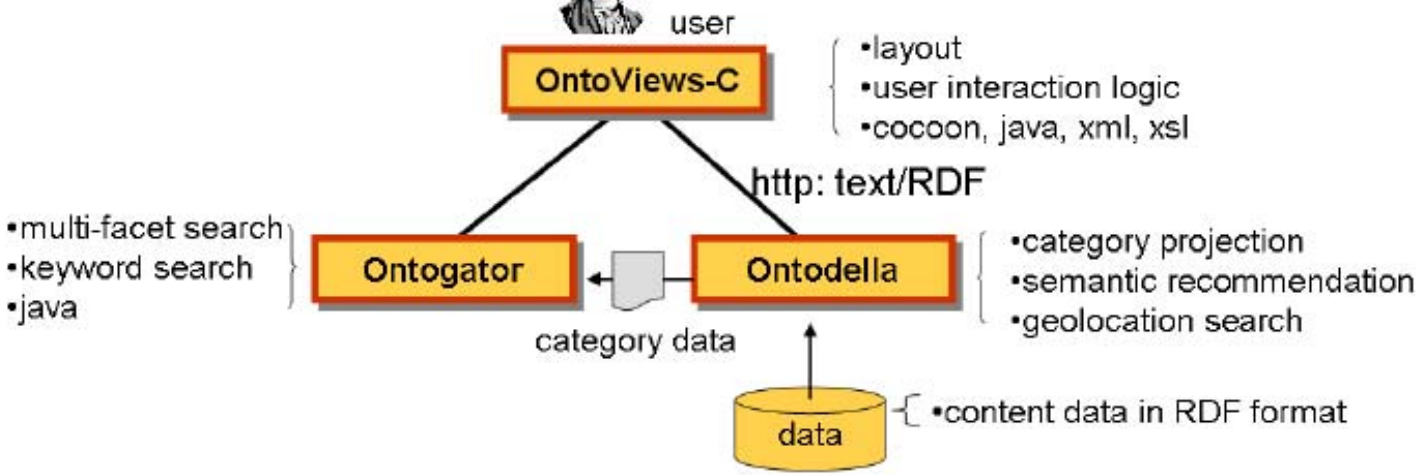
- Providing alternative views to eGov link library content
- Aggregating relevant content automatically from different



(Sidoroff & Hyvönen, ISWC 2005 WS)

OntoViews Framework

- Ontogator + Ontodella + Apache Cocoon
- (Makelä, Hyvönen, Viljanen, ISWC 2004)
  (Mäkelä, Hyvönen, Saarela, ISWC 2006)
  (Viljanen; Hyvönen, Känsälä, DEXA WS, 2006)



Semantic Faceted Search Engine ONTOGATOR

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Integrates view-based search paradigm with semantic web
  - Ontologies and reasoning
- Used to build 8 different systems
- Scales up to millions of search items and hundreds of thousands of categories
- (Mäkelä, Hyvönen, Saarela, ISWC 2006)

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Semantic Recommendation Server ONTODELLA

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

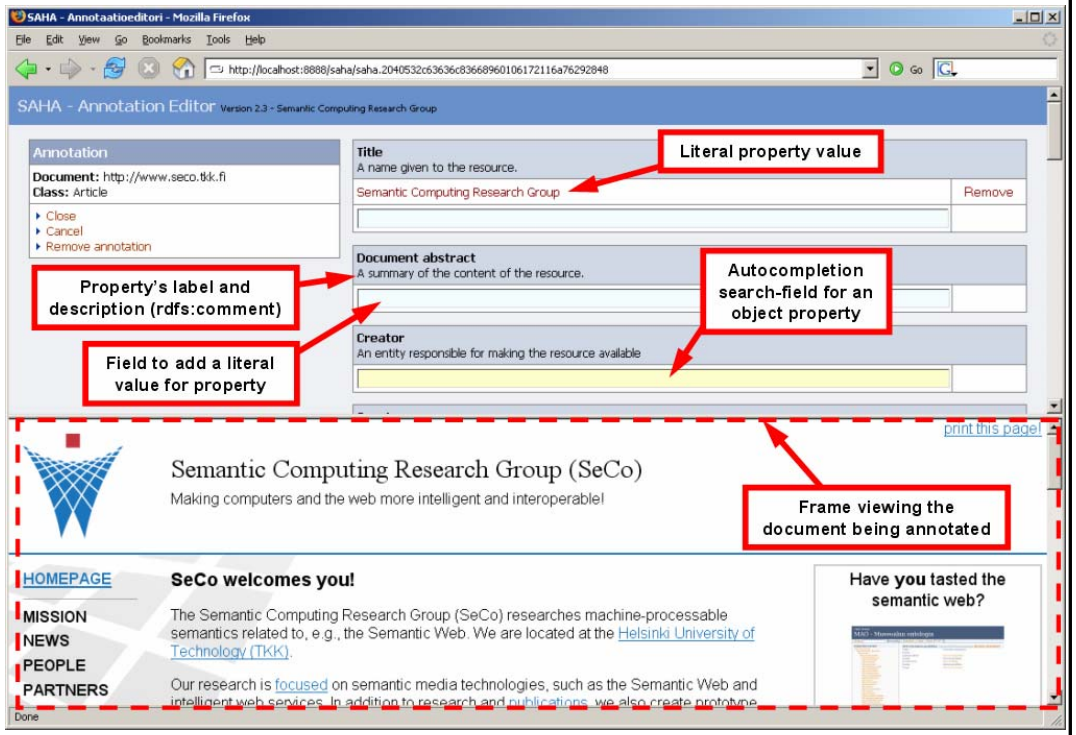
- For creating semantic recommendations between resources
- Provides
  - logic-based services for projecting facets
  - logic-based recommendations with explanations
- (Viljanen, Hyvönen, Känsälä, DEXA 2006 WS)

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Annotation tool SAHA

- Adapts to different annotation schemas
- Support distributed annotation work and ontology population
- Easy to use with a web browser

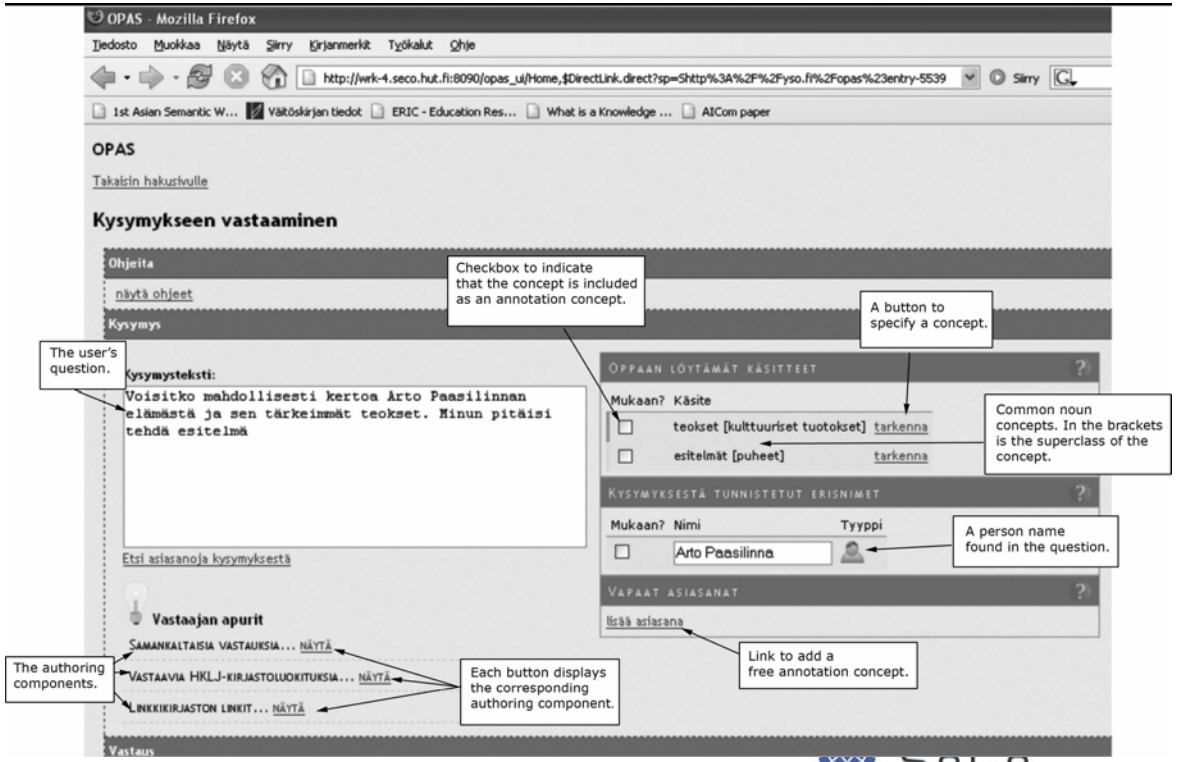


(Valkeapää, Hyvönen, ISWC 2006 WS)
(Valkeapää, Alm, Hyvönen, JUCS, 2007)

Semi-automatic semantic annotation and authoring tool OPAS

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Support help-desk question-answering (QA)
  - Annotating QA-pairs and helping authoring answers
- National ”Ask the librarian” service as a case study



(Vehviläinen, Hyvönen, Alm, ISWC 2006 WS)

UNIVERSITY OF HELSINKI

Using and visualizing historical geospatial information

Figure 6: A searchable map interlinked with the semantic portal MuseumFinland.

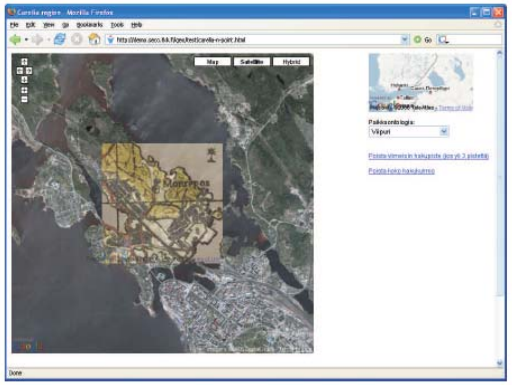
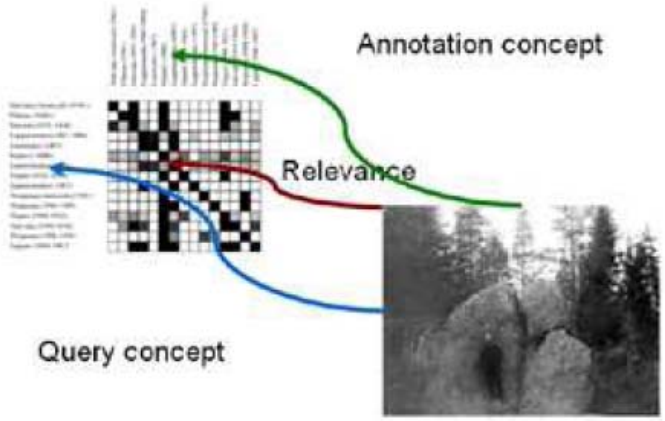


Figure 4: Using multiple maps simultaneously. A historical Karelian map depicting the park of Monrepos in Viipuri is shown semi-transparently on top of a modern satellite image provided by the Google Maps service.

Figure 5: Search results using the n-point search: Viipuri, Koivisto and Makslahti are matched.



(Kauppinen et al., FAIS 2006)

Figure 3: Annotation and indexing concepts matched.

Other topics

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Uncertainty in ontologies
  - Representing and reasoning with spatial uncertainty
    - » (Holi, Hyvönen, Springer 2006)
  - Fuzzy view-based search
    - » (Holi, Hyvönen, ASWC 2006)
- Representing and reasoning with historical geospatial changes
  - An ontological model of Finnish communes and counties
    - » (Kauppinen, Hyvönen, Springer 2006)
- Semantic autocompletion
  - » (Hyvönen, Makelä, ASWC 2006)
- Combining text- and ontology-based searching
- Automatic annotation using NLP

UNIVERSITY OF HELSINKI

SeCo
semantic computing

What Next?

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Semantic Web 2.0 – Intelligent Collaborative Services
  - Idea: synergy of Web 2.0 and Semantic Web
  - Based on FinnONTO infrastructure
  - 3 years, 39 funding organizations, 1M€ / year
  - 2008-2010, if funding application is approved

UNIVERSITY OF HELSINKI

SeCo
semantic computing

Conclusions

HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Semantic web is coming
- An ontology-based infrastructure is needed for it
- Open infrastructure enables development of practical applications
- FinnONTO is an experiment of this on a national Finnish level

- Papers, software, ontologies, demos, and pilot systems available at:
  - http://www.seco.tkk.fi/

- Thank you
- Questions?

UNIVERSITY OF HELSINKI

SeCo
semantic computing