

Building a National Semantic Web Ontology and Ontology Service Infrastructure —The FinnONTO Approach

<u>Eero Hyvönen, Kim Viljanen, Jouni Tuominen, and Katri Seppälä</u>

Semantic Computing Research Group (SeCo) Helsinki University of Technology (TKK) and University of Helsinki Finland

http://www.seco.tkk.fi/firstname.lastname@tkk.fi







- FinnONTO A Vision of a National Semantic Web Infrastructure
- Major Infrastructure Components
 - System of Mutually Aligned Cross-domain Ontologies
 - Public Ontology Services
- Usage Tested/Demonstrated in FinnONTO Applications
- Next: Usage in Legacy Systems Using the ONKI Living Lab







FinnONTO – A Vision of a National Semantic Web Infrastructure





FinnONTO Thesis



- 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
- Especially useful in
 - Cross-domain applications
 - Collaborative Web 2.0 applications





FinnONTO Infrastructure Components



HELSINKI UNIVERSITY OF TECHNOLOGY

Media Technology

1. System of mutually aligned cross-domain ontologies

- Distributed collaborative ontology development
- Centralized global view for end-users

2. Public ontology services

- Centralized ontology Services for legacy applications
- For mashups and web services
- 3. Metadata schemas
 - Cultural heritage, health, learning, geodata, ...
- 4. Tools
 - Metadata creation / annotation
 - Semantic portal creation
 - » Search and browsing



Industrial & Public Organization Consortium



HELSINKI UNIVERSITY OF TECHNOLOGY

Media Technology

- FinnONTO: National Semantic Web Ontology Project in Finland
 - -2003-2004
 - » 14 funding organizations
 - -2004-2005
 - » 16 funding organizations
 - 2005-2006
 - » 30 funding organizations
 - -2006-2007
 - » 37 funding organizations
- FinnONTO 2.0: Semantic Web 2.0 Intelligent Collaborative Services
 - -2008-2009
 - » 38 funding organizations
 - _ 2009-2010



















KANSALLISKIRJASTO



Kansanterveyslaitos Folkhälsoinstitutet National Public Health Institute



UNIVERSITY OF HELSINK Viikin tiedekiriasto



m-cult















SUOMEN MAATALOUSMUSEO SARKA

Suomalaisen Kirjallisuuden Seura



Valtiovarainministeriö Finansministeriet Ministry of Finance



TEKNILLINEN KORKEAKOULU Viestintätekniikka



Suomen valokuvataiteen museo Finlands fotografiska museum The Finnish Museum of Photography





TEKNILLINEN KORKEAKOULU Kartografia ja geoinformatiikka



















HELSINGIN KAUPUNGINKIRJASTO HELSINGFORS STADSBIBLIOTEK HELSINKI CITY LIBRARY



System of Mutually Aligned Cross-domain Ontologies





The Way to Go



- Started a national cross-domain ontologization process
 - Making contents of different domains 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 Ontology Development



- Class ontologies
 - Keywords and concepts used in indexing nationally
 - Population strategy: mainly Thesauri -> Ontologies
- Instance ontologies
 - Persons, places, ...
 - Populated mainly from databases





A Method for Transforming Thesauri into Ontologies



- Input: a thesaurus using standard ISO [2788] / SFS 5471
 - Semantic relations: NT, BT, RT, USE, USED FOR, ...
 - Used widely in Finland
 - Fairly large vocabularies, eg. YSA 23,000 terms
- Output: a light-weight ontology
 - Equivalence relationship
 - » Semantic disambiquation of term meanings
 - » Preferred vs. non-preferred terms (USE / USED FOR)
 - » Multilinguality
 - Hierarchical relationship
 - » Based on NT / BT
 - » Transitive subclass hierarchy sceletons
 - » Part-of relations distinguished from subclass-of
 - Associative relationship
 - » Based on refining RT relations of the thesaurus







General Principles Used

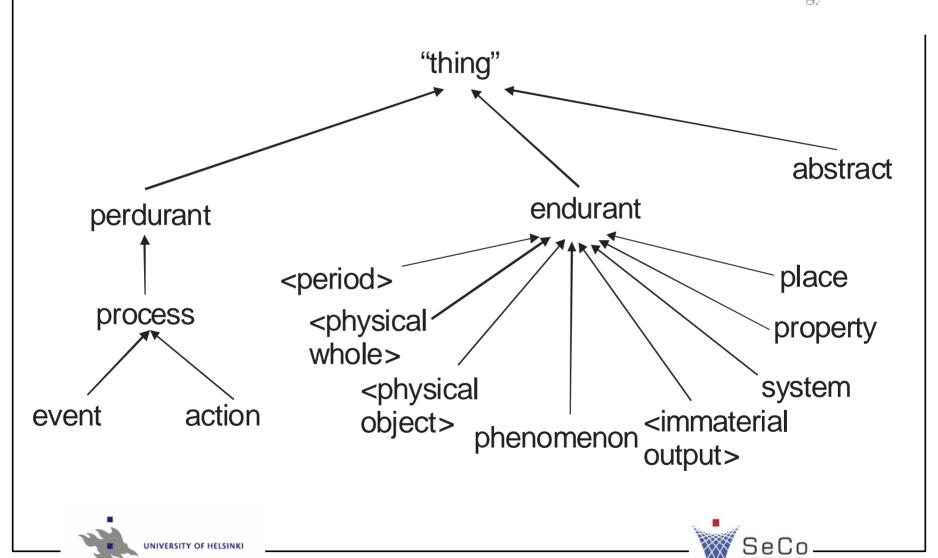
- Minimal useful ontologization but more than mechanical SKOS transformation
 - » Thesaurus semantics -> basic RDFS semantics
 - » Applications add application specific properties etc.
- Main category facets based on DOLCE
 - » Abstract (quantities, ...)
 - » Endurant (enduring objects and things)
 - » Perdurant (events, processes)
 - To be used for event-based annotations





YSO - The Finnish General Upper Ontology (Upper Structure)







- Minimize multiple inheritance
 - » Allowed only within major facets
- The ontology is a model of the world not of the thesaurus
 - » Cf. SKOS where terms are instances of skos:Concept
- Terms are represented systematically as classes
 - » Instance vs. class distinction is left to applications
 - E.g. Halley's comet, Peace of Paris, ...
 - » Free indexing keywords -> instantiate a class
 - E.g. flower types, animal species, ...

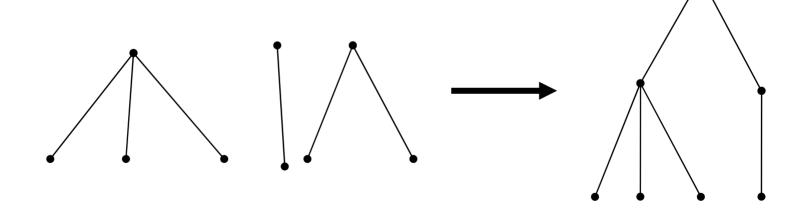




Steps



- Add missing links in the subclass hierarchy
 - E.g. the Cultural thesaurus MASA (6000 concepts) had 2500 top concepts

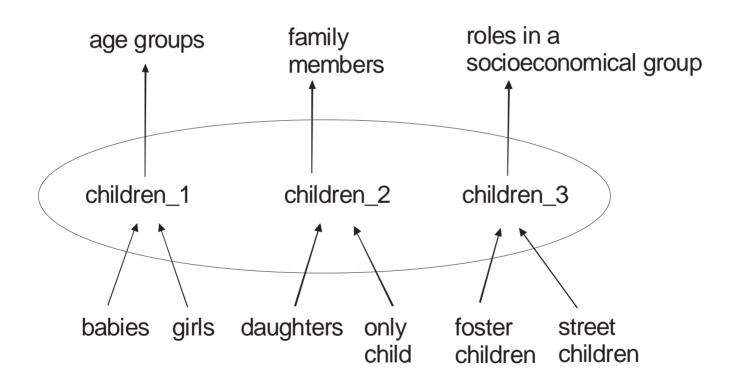








- Disambiguate term meanings
 - E.g. child / children

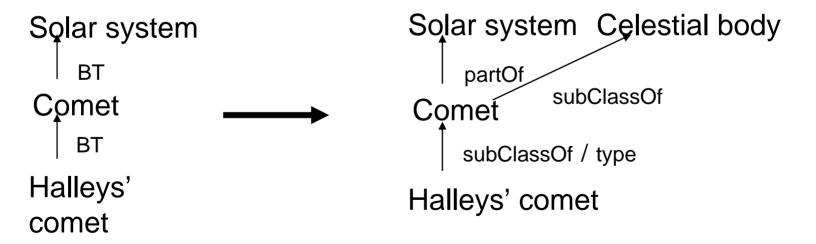








- Disabiguate BT relation meanings
 - subclassOf vs. part-of vs. instantiation
 - Halleys comet BT solar system (part-of)
 - lion BT cat animals (subclass of)









Check transitivity of the BT chains

make-up mirror BT mirror mirror BT furniture

=> make-up mirror subClassOf furniture?

furniture

BT

mirror

BT

BT

make-up mirror





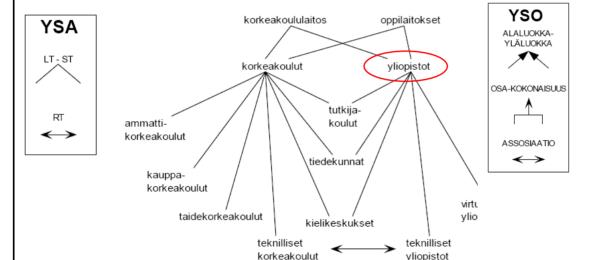
Example: University Terminology YSA Thesaurus -> YSO Ontology



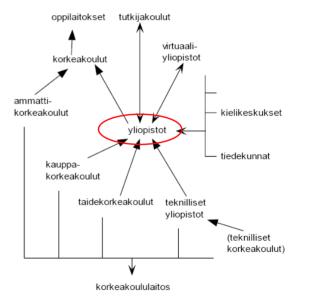
HELSINKI UNIVERSITY OF TECHNOLOGY

Media Technology

Thesaurus



Ontology









- Concept classes
 - Used for annotations
 - E.g. cat, unemployment, ...
 - Can be used by instantiation
 - » Free indexing terms
 - E.g. plant types, places, etc.
 - Can be used by instatiating nearest superclass
- Grouping concepts (collections)
 - For organizing concepts, not for annotations
 - E.g. endurant, clothes-by-sex
- Aggregate concepts
 - Aggregate multiple meanings
 - » E.g. child (age + family relation + econosocial role)
 - Desired for
 - » simplifying annotations
 - » compatibility with thesauri terms





Other FinnONTO Class Ontologies Using SKOS



- Using existing classifications and vocabularies as they are
 - Mechanical tranformation into SKOS
 - Accepting existing semantic ambiguities and confusions
 - Finnish translations produced as add-on
- Examples
 - ICONCLASS
 - Medical Subject Headings (MeSH)
 - Library classifications (e.g. HKLJ based on Dewey Decimal)







- Places
 - Small number of classes (< 1000)
 - Lots of particular places (N x 1 000 000)
- Actors
 - Small number of classes (< 1000)
 - Persons, groups, organizations (N x 100 000)
- (Historical) events



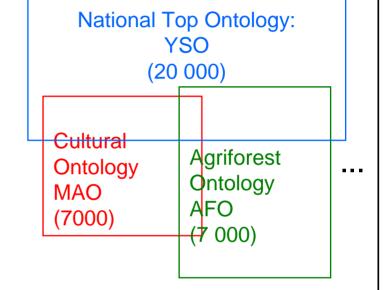


FinnONTO National Ontologies



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

- Top ontology YSO as semantic glue
- Merges overlapping domain ontologies
 - Cultural ontology MAO
 - Geo-ontology SUO
 - Spatiotemporal geo-ontology SAPO
 - Actor ontology TOIMO
 - Applied arts ontology TAO
 - Photography ontology VALO
 - Agriforest ontology AFO
 - Ontology of Finnish History HISTO
 - ...



- Alignments with other classification systems and top ontologies
 - » HKLJ + YSO

(library domain)

» ICONCLASS + YSO

(fine arts domain)

» MeSH + YSO

(medical domain)

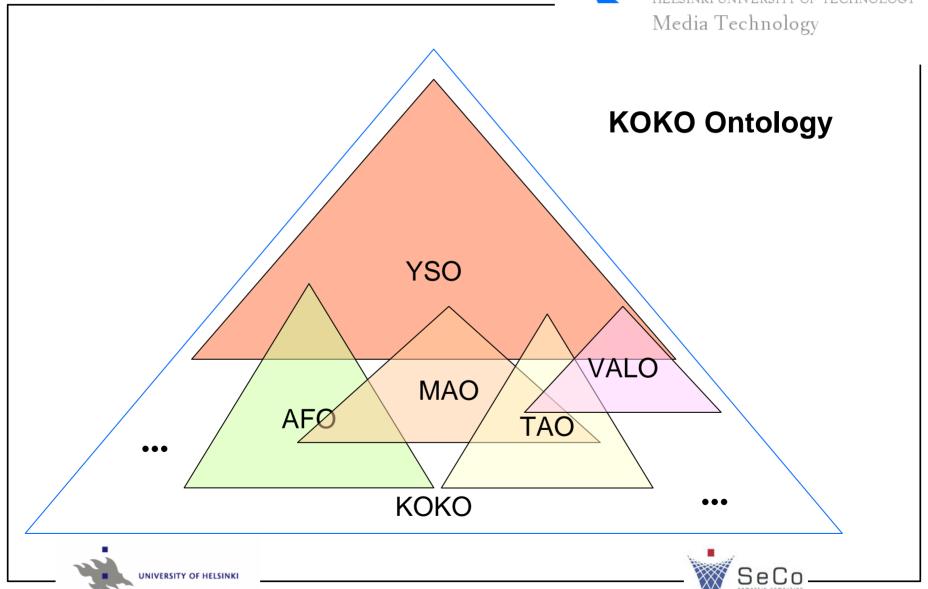




KOKO: Finnish Collaborative Holistic Ontology: Ontology developer groups' view



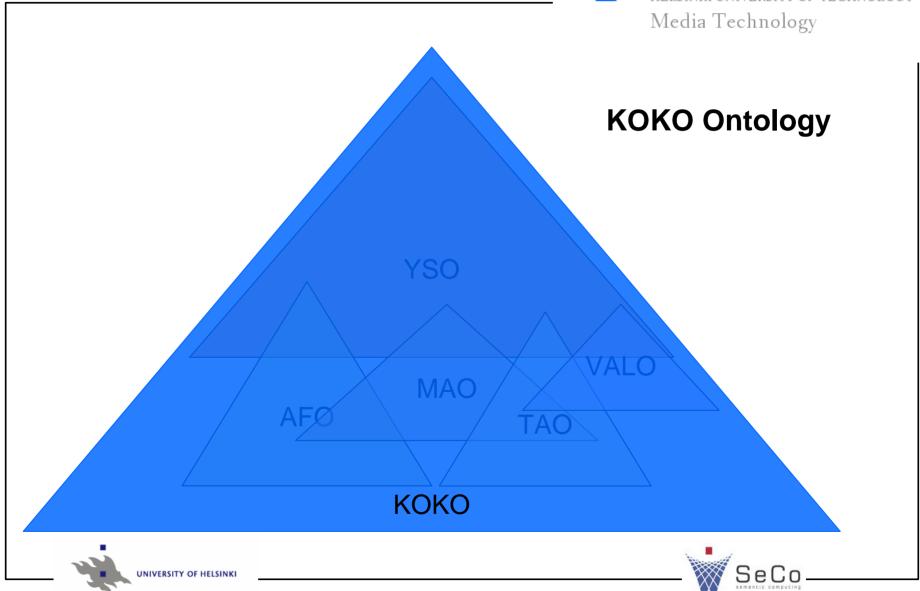
HELSINKI UNIVERSITY OF TECHNOLOGY



Finnish Collaborative Holistic Ontology KOKO: Ontology end-users' view



HELSINKI UNIVERSITY OF TECHNOLOGY



KOKO National Collaboration Principles



- Start thinging global / cross-domain
 - Every domain should consider also cross-domain reuse of concepts
- Add machine semantics
 - Proceed in small steps
 - Keeps e.g. funding agencies happy
- Reuse existing vocabulary work and datasets
 - Maintain interoperability with the past and with others
- Establish collaboration network of independent domain expert groups
 - Based on current thesaurus developer groups
- Coordinate nation wide the work by a steering group
 - Now by the FinnONTO research project
 - Later by the National Library
- Respect different ontological views
 - Encourage reuse of others' ontology structures but do not force
 - It is not possible to come up with only one ontology view
 - » Different application views and opinions will survive





Ontology Production: Protege-editor for Aligning Ontologies



HELSINKI UNIVERSITY OF TECHNOLOGY

Media Technology 🔜 maotaoyso_txt.rdf-xml.rdf-xml. Protégé 3.2 (file:\D:\kaseppal_cvsroot\ontologies\varmuuskopio_ks\svn_checkout\mao\maotaoyso_txt.rdf-xml.rdf-xml.pprj, OWL / RDF Files) File Edit Project OWL Code Tools Window Help 🗖 protéaé ● Metadata (mao) ● OWLClasses ■ Properties ◆ Individuals ■ Forms SUBCLASS EXPLORER For Project: maotaoyso txt.rdf-xml.rdf-xml For Class: 6 kasityot (instance of MaoConcept)

Inferred View 💣 醇 🍖 💂 Annotations Asserted Hierarchy Value Lang e kartat = vso kartat rdfs:comment käsin tai käsin koneella eri materiaaleista valmistetut esineet kartunta ▶ ■ kasetit = yso kasetit = tao:kasetit kasikayttoinen ■ kasiteollisuus = tao:käsiteollisuus = vso käsiteollisuus or or 😞 🐟 Asserted Conditions ■ kasityokasvatus = yso käsityökasvatus = tao:käsityökasvatus NECESSARY & SUFFICIENT pyso käsityöt 2 hoyhentyot NECESSARY mao käsitellyt ⊑ koulukasityot ▶ □ kuitutvot = tao:kuitutvöt □ lakkatyot = tao:lakkatyöt = yso lakkatyöt lasityot = yso lasityöt ► • luutvot = tao:luutvõt = vso luutvõt ▶ ■ metallityot = tao:metallityöt = yso metallityöt nahka- ja turkistyot 💣 🛖 🍖 🐴 🍖 Disjoints puhdetyot pullolaivat ▶ ■ punontatyot = tao:punontatyöt = yso_punontatyöt ▶ **⑤** puutyot ≡ tao:puutyöt ≡ yso puutyöt sulkatyot teelmat tekstiilitvot skos:broaderGeneric e kaste = tao:kaste = vso kaste e kastealtaat = tao:kastealtaat = yso kastealtaat e kasvihuoneet ≡ yso_kasvihuoneet kasvit = yso_kasvit = tao:kasvit altl abel e kattaukset = tao:kattaus = yso kattaus Value e katteet = yso katteet = tao:katteet katukivet katupeilit e kaukoputket = vso kaukoputket ▶ ⑤ kauppa ≡ yso kauppa Value kasitvot ● Logic View ○ Properties View Protege.exe Microsoft PowerPoint FI (10:09 🎒 start (S) ONKI-Selain : Ontolo. Inbox for katri.seppal... maotaoyso_txt.rdf-x.. UNIVERSITY OF HELSINKI

Checking Ontologies Against Each Other Using the ONKI Browser



HELSINKI UNIVERSITY OF TECHNOLOGY

Media Technology





Public Ontology Services





Centralized Ontology Services



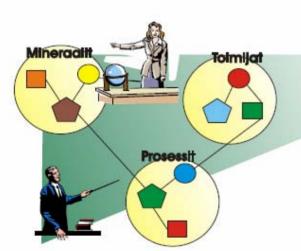
HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

2. Information Searchers

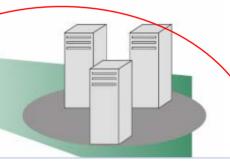
- Support concept-based search
- Keyword disambiguation
- Finding the right search concepts



Nokia: company or city?







2. Information Indexers

- Support indexing concept finding
- Keyword disambiguation
- Support indexing patterns



1. Ontology Developers

- Colloborative development of interdependent ontologies

- Versioning and support for updates



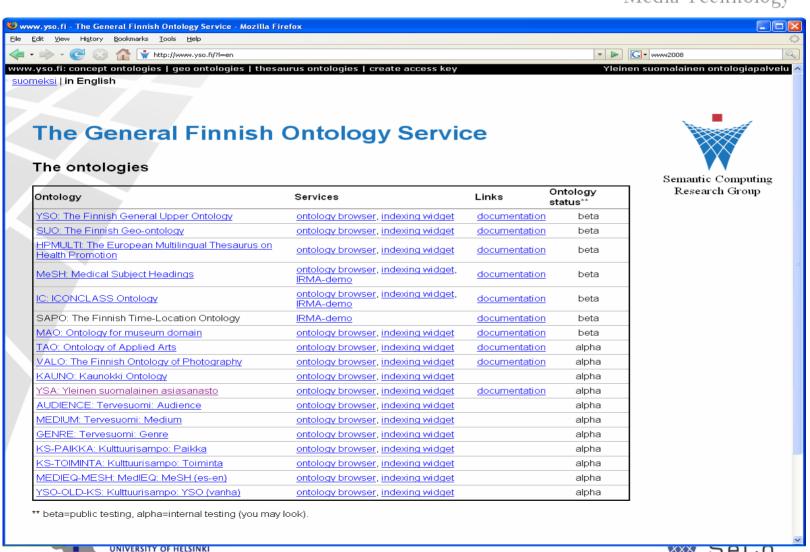
National Ontology Service YSO:

http://www.yso.fi



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

Seri

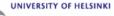


Utilizing ONKI Ontology Services as Mashups in Legacy Systems



- Idea related to Google Maps
 - Provide centralized service with end-user interfaces
 - Use it to create mashup services based on AJAX
 - Also Web Services can be used
- ONKI provides services for
 - Concept finding (semantic autocompletion)
 - Semantic disambiguation
 - Concept selection, browsing, changing
 - Also maps are used (based of Google Maps)
 - » For the General Finnish Location Ontology
- Services are used in legacy system using the ONKI Widget
 - Ready-to-use mashup interface component
 - 1 line of Javascript on your HTML page!





ONKI Ontology Servers



- ONKI SKOS
 - Input your SKOS vocabulary and the services are ready to use on the web!
 - Supports also simple RDFS-ontologies
- ONKI Geo
 - For publishing geo-ontologies with Google Maps services
- ONKI Instance
 - For publishing instance ontologies (eg. authorities)





ONKI SKOS for Vocabularies & Simple Ontologies

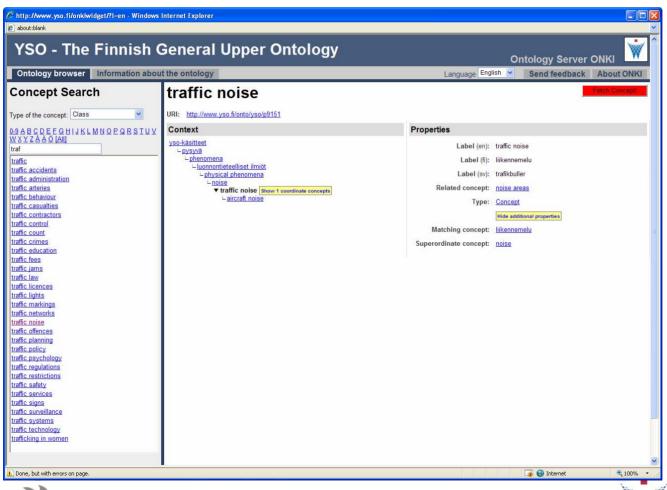
E.g. http://www.yso.fi/onki/yso/?l=en

UNIVERSITY OF HELSINKI



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

Supports W3C SKOS & simple RDF(S) ontologies



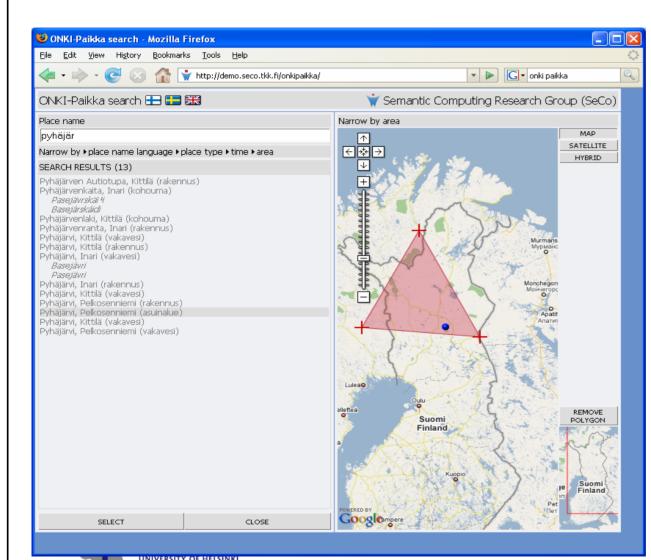


ONKI Geo for Geo-ontologies

http://demo.seco.tkk.fi/onkipaikka/



Media Technology



800,000 contempory Finnish places

Finnish historical places

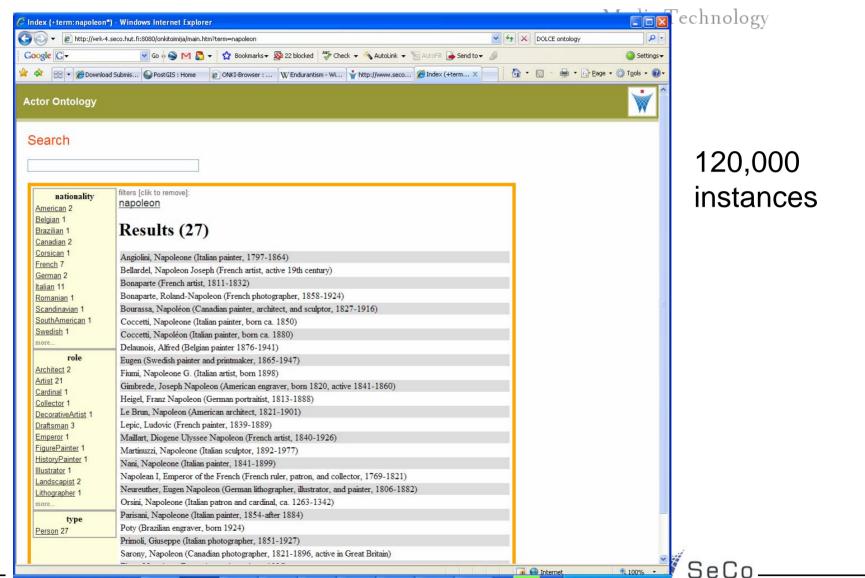
Millions of places abroad (GNS,GNIS)



ONKI Instance Publishing Getty ULAN Register



HELSINKI UNIVERSITY OF TECHNOLOGY

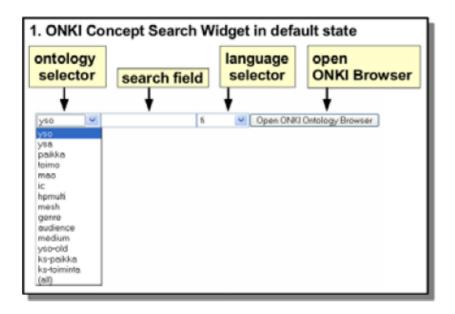


ONKI Widget for Mashups

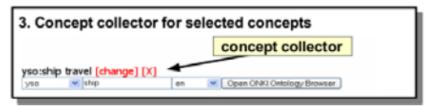


HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

- Ontology services are automatically available after publishing a vocabulary or ontology with ONKI
- Simple AJAX-based widget for creating mash-ups
 - Direct Web Remoting (DWR) API





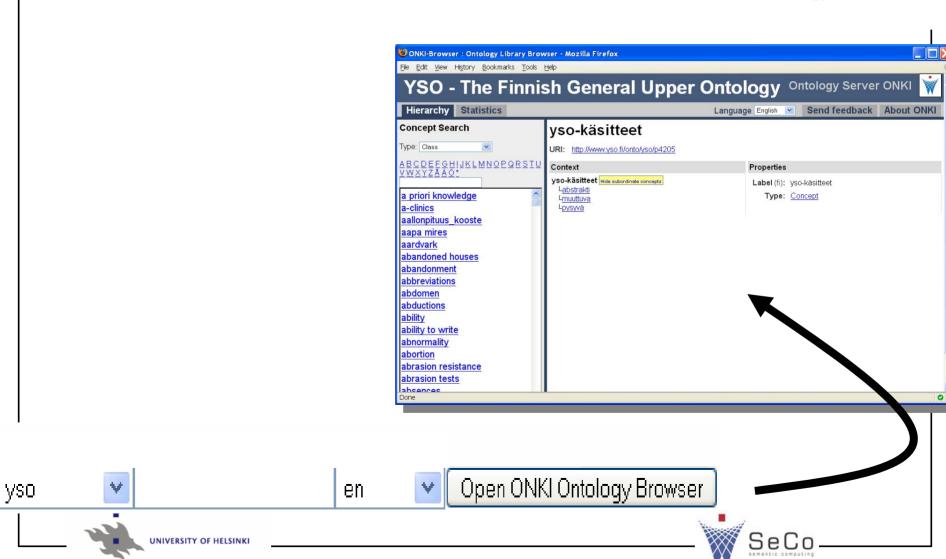








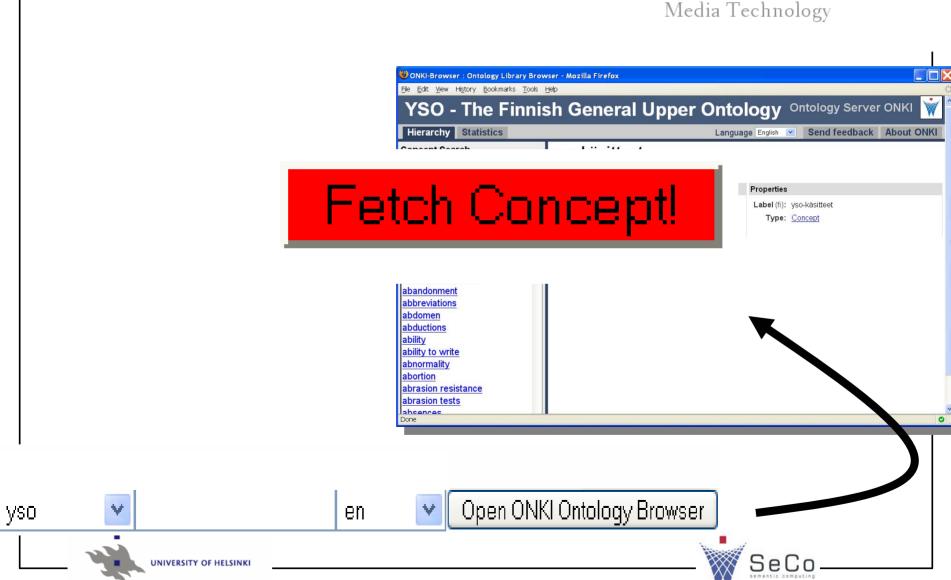
Media Technology





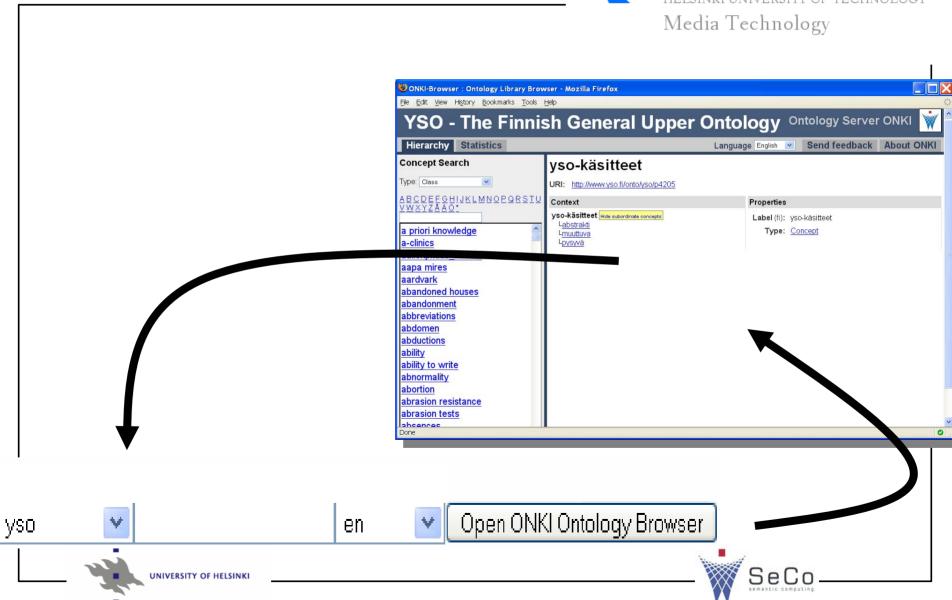
HELSINKI UNIVERSITY OF TECHNOLOGY

Media Technology

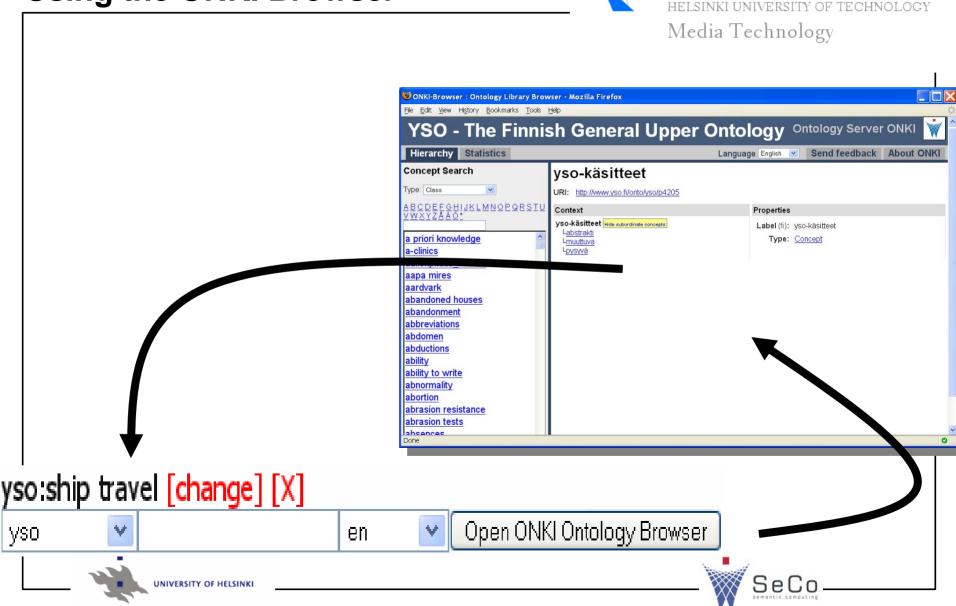




HELSINKI UNIVERSITY OF TECHNOLOGY







ONKI Widget JavaScript Code



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

To the HTML HEAD section:

```
<script type="text/javascript"
src="http://www.yso.fi/onki.js"></script>
```

Update an HTML Form input field to ONKI Widget field:

```
<input id="dc:subject"
  onkeyup="onki['yso'].search()"/>
```

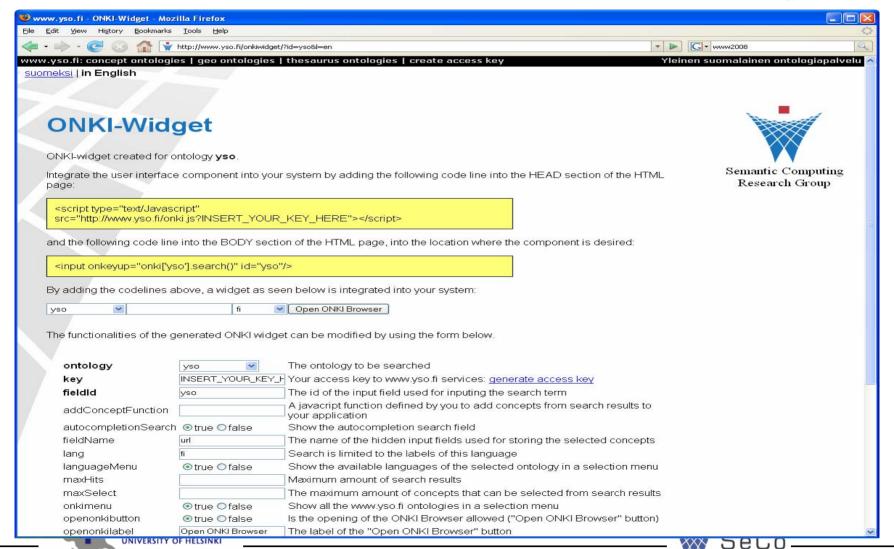




ONKI-Widget Generator: specify->copy->paste into your web page http://www.yso.fi/onkiwidget/?l=en



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology



Using ONKI Widget in a Legacy System:

http://www.yso.fi/onki/yso/app/annotation/



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

Secondario computing

🐸 Annotation demo using ONKI Mash-Up Component and IRMA Context Browser - Mozilla Firefox		
<u>F</u> ile <u>E</u> dit <u>V</u> iew I	Higtory <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp	
@		▼ ▶ G ▼ www2008
The ONKI Mash-Up component in the "Subject" field is integrated into the system by adding only two lines of code.		
The component can also be used for annotating with terms from traditional thesaurus. Thus it's possible to integrate the component to a legacy system, which isn't aware of URIs.		
In addition to using ONKI Mash-Up component, the <mark>IRMA</mark> widget can be used to visualize the hierarchy of concepts (see demo part 3). IRMA is attachable to any HTML-based annotation system with JavaScript.		
Title	Toy truck	Text field for the title of the document.
	Orange toy truck with blue wheels.	
Description		Text field for the description of the item.
Picture	Add new picture	Field for the picture of the item
Demo part 1: fetching ontological concepts		
Subject	yso:car makes [change] [X] yso	Semantic autocompletion search field for fetching concepts/instances from ontologies. Type some text in the text field to perform a search. By clicking a search result the concept's URI and label are fetched. The ontology used in search can be changed by using the leftmost menu. The language used in search can be changed by using the rightmost menu. The selected ontology can be opened in ONKI Browser by pressing the "Open ONKI Browser" button. In browser the "fetch concept" button can be pressed to fetch the selected concept to the cataloguing system. The label of a fetched concept functions as a link to the HTML representation of the concept. A fetched concept can be changed to another concept by clicking the "change" link next to the concept's label. A fetched concept can be removed by clicking the X-symbol next to the concept's label.



Selected FinnONTO Applications





Selected Collaborative National Semantic Portals



- MuseumFinland Finnish Museums on Semantic Web
 - http://www.museosuomi.fi
- CultureSampo Finnish Culture on the Semantic Web
 - http://www.kulttuurisampo.fi
- HealthFinland Finnish Health Information on the Semantic Web
 - http://www.seco.tkk.fi/applications/tervesuomi/
 - http://www.tervesuomi.fi
- ...







- Official publication event of the National ONKI Living Lab
 - http://www.yso.fi (already there and operational)
 - Distribute open source ontologies
 - Support Public ONKI Services
 - Support publishing other groups ontologies
 - » Let us know if you would like to try it out!
 - Gain legacy end-user experiences of using the technology
- Continue aligning Finnish vocabularies and ontologies with those in other languages







- Semantic Web needs cross-domain content infrastructure
 - Shared ontologies
 - » Collaborative development needed
 - Ontology services for applications
 - » Mashups and web services
- FinnONTO tries to facilitate this on a national level in Finland
- Technology being tested in several collaborative semantic portals
 - MuseumFinland, CultureSampo, HealthFinland ...
- Next step: ontology support for lecagy systems using the ONKI Living Lab



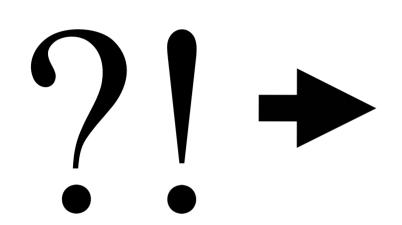


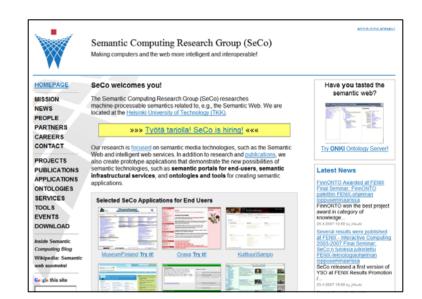
Thanks!

More info, publications & demos on the SeCo-web



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology





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



