



Kohti suomalaisen semanttista webiä -- ontologioita, palveluita ja sovelluksia

Prof. Eero Hyvönen
TKK Viestintätekniikka
Helsingin yliopisto, tietojenkäsittelytieteen
laitos

Semanttisen laskennan tutkimusryhmä

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



Symposion avaus



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Tervetuloa TKK:uun
- Pahoittelemme salin pienuutta
 - Suurempaa ei ollut saatavilla
 - Näin laaja kiinnostus aihetta kohtaan (330 ilmoittautunutta) yllätti
- Symposion tarkoitus
 - Esitellä kansallisen FinnONTO-hankkeen visioita ja toimintaa
 - » Tähänastiset tulokset 2003-2005
 - » klo 9:00-15:30
 - Keskustella jatkotoimenpiteistä 2006-2007
 - » klo 15:30-16:30
 - Semanttisesta webistä kiinnostuneiden verkottuminen
 - » Osa Fenix-ohjelman Tietämyksen hallinta ja semanttinen web –aiheryhmän toimintaa
 - » Esim. uusien hankeideoiden kehittäminen Tekesille

Käytännön järjestelyitä



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Ohjelma, osallistujaluettelo ja yhteenveto FinnONTO-hankkeesta saatavilla
- Kahvitukset
 - klo 10:30 ja 15:00
 - Kiitos tuesta Tekesin Fenix-ohjelmalle
- Lounas (omakustanteinen)
 - klo 12:30-13:30
 - Esimerkiksi aulan yhteydessä on ravintola



- Mikä on semanttinen web?
- Miksi tarvitaan kansallisen semanttisen webin infrastruktuuri?
 - Ongelma
 - Teesit ratkaisumallille
- Suomalaiset semanttisen webin ontologiat (FinnONTO)-projekti
 - Tavoitteet
 - Tuloksia

Mikä on semanttinen web?



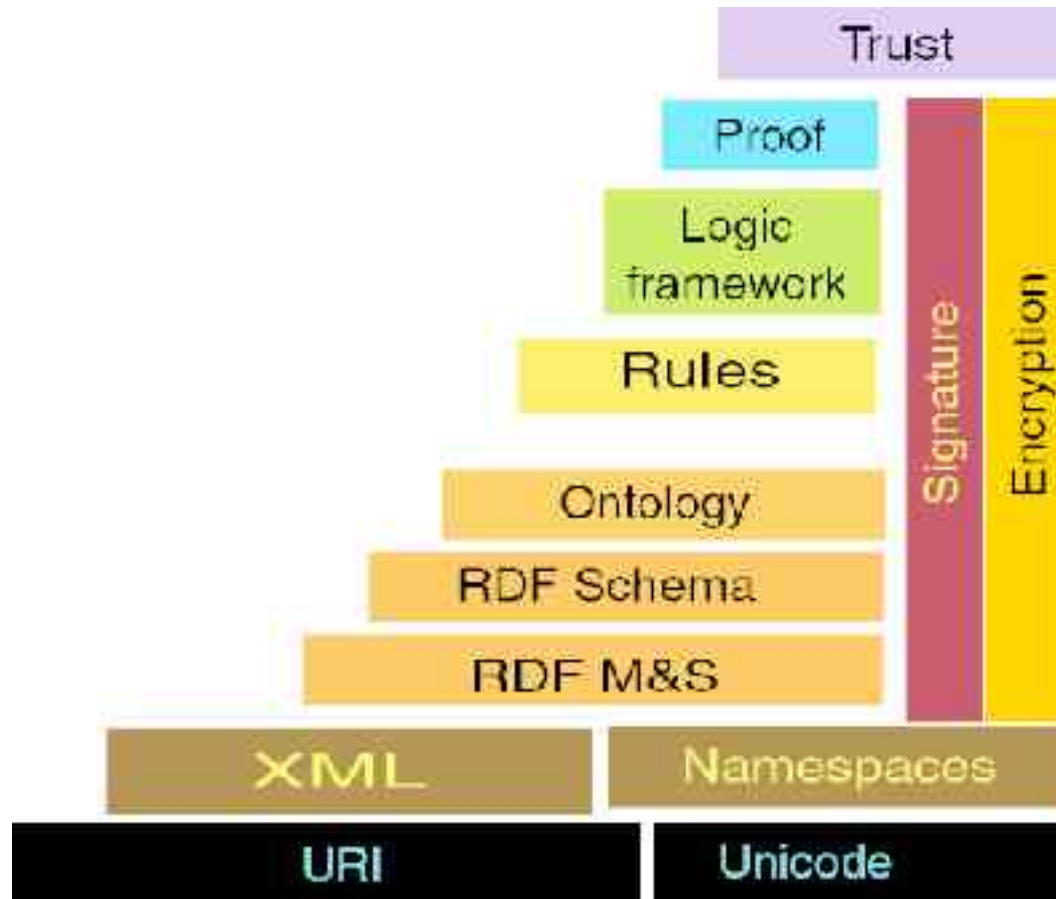
- Webin metadatakerros, jossa webin sisältö ja palvelut on kuvattu tietokoneelle sen ymmärtämällä käsitteistöillä, ontologioilla
- Kansainvälinen tahtotila webin tulevaisuudesta
 - World Wide Web Consortium (W3C) Activities
 - » Semantic web activity, Web service activity
- 2000-luvulla muodostunut uusi oma tutkimusalue
 - Web-teknologioiden ja tietämystekniikan yhdistelmä
 - Omat konferenssisarjat
 - » ESWC, ISWC, ASWC, ...
 - Omat tieteelliset jurnaalit
 - » Journal of Web Semantics, Int. Journal on Semantic Web and Information Systems, Int. Journal of Metadata, Semantics and Ontologies,...



Semanttisen webin "teknologiakakku-malli"



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology



W3C[®]

(Tim Berners-Lee)



UNIVERSITY OF HELSINKI

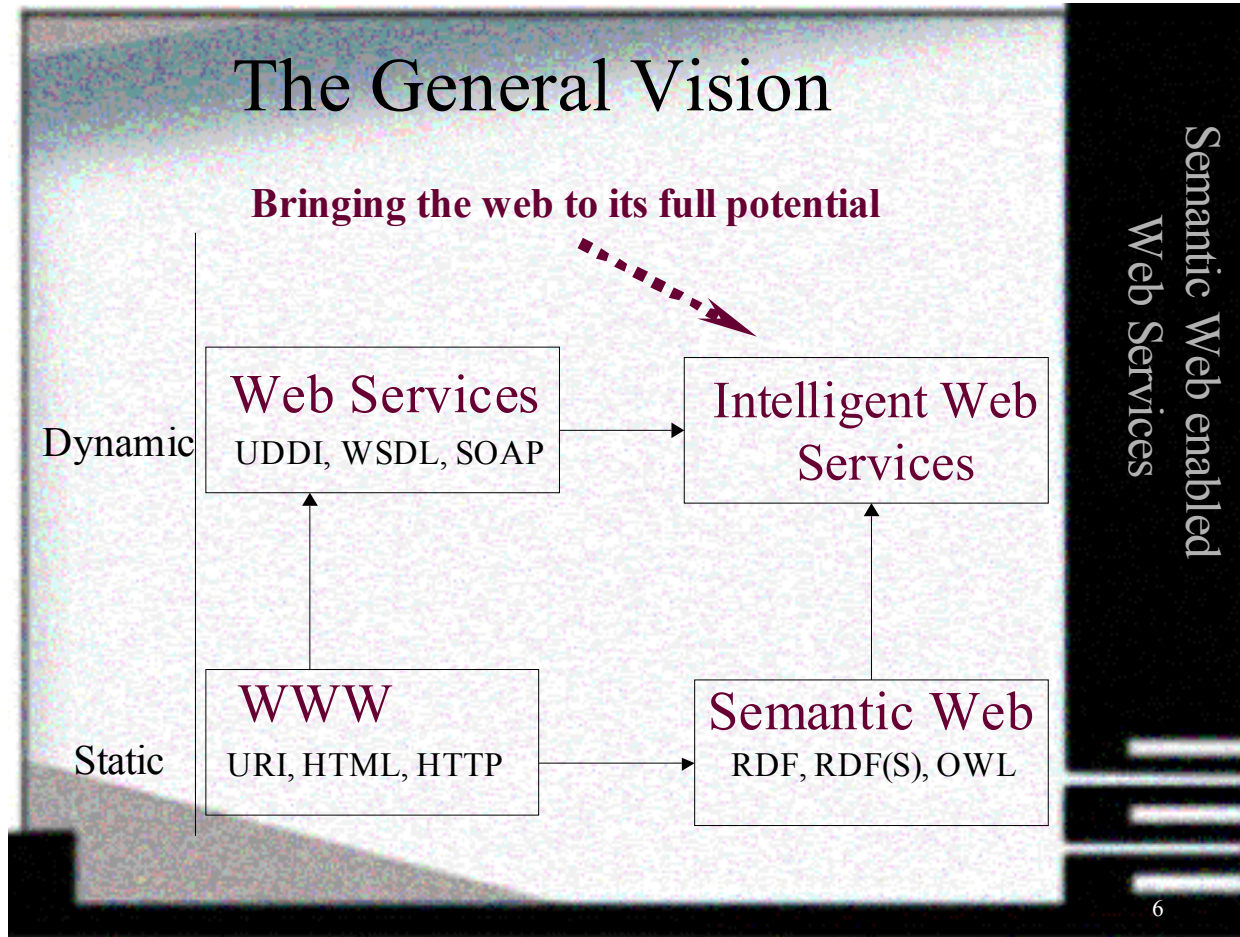


UNIVERSITY OF HELSINKI



SeCo

Webin semanttiset megatrendit



(Dieter Fensel, 2002)



- Älykkäämmät web-sovellukset ihmisille
- Yhteentoimivat järjestelmät (interoperability)
 - Sisällöt yhteismitallisiksi merkitystasolla
 - Järjestelmät kykenevät kommunikoidaan keskenään
 - Olemassaolevien järjestelmien uusiokäyttö webissä
- Kustannustehokkuus
 - Sisältöjen ja käsitteistöjen jakaminen
 - Web-palveluiden hajautettu ”julkaiseminen”
 - Toimintojen koostaminen laajemmiksi web-palveluiksi

Kansallinen ongelma



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Neljä faktaa
 - Semanttinen web = webin seuraava sukupolvi/kerros
 - Ontologiat = semanttisen webin ”silver bullet”
 - Suomalaiset ontologiat = ei ole
 - Tarttis tehrä jotakin
- Ilman yhteisiä ontologoita
 - aitoa semanttisuutta ei webiin saada,
 - semanttinen web pirstaloituu yhteentoimimattomiksi saarekkeiksi,
 - sisältötyötä tehdään turhaan moneen kertaan ja
 - web-sovellusten kehittäminen tulee vaikeammaksi ja kalliimmaksi.

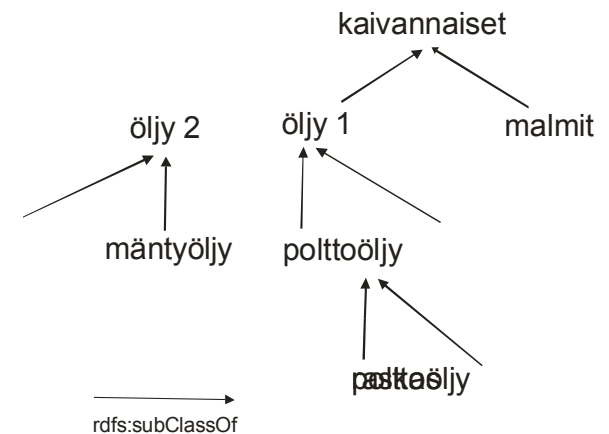


FinnONTO-hankkeen teesit (1)



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Semanttisen web tarvitsee julkisen sisältöinfrastruktuurin
 - kuten liikenne tarvitsee tieverkoston
 - kuten energiahuolto tarvitsee sähköverkon ja standardit
 - kuten tietoliikenne tarvitsee GSM tai 3G-verkon
- Infrastruktuurin ytimenä ovat
 - laajat **referenssiontologiat**
 - näihin liittyvät keskeiset **ontologiapalvelut**
 - kansainväliset **standardit**
 - **työkalut**

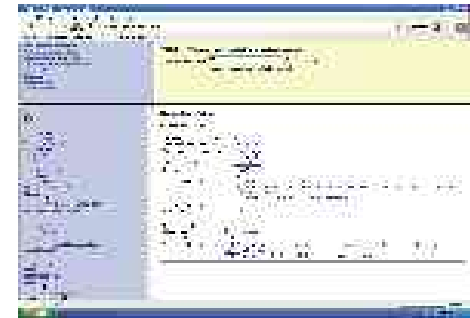


FinnONTO-hankkeen teesit (1)



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Asiasanastoista on siirryttävä ontologioihin
 - Ihmisen ymmärtämistä sanoista koneen ”ymmärtämiin” merkityksiin
- Kansallinen ontologisointiprosessi tarpeen
 - Eri alojen/organisaatioiden sisältöjen yhdistämiseksi
 - Ydinontologioita ylläpidetään julkisesti
 - » Laaja sitoutuminen välttämätöntä
 - Ydinontologiat julkaistaan open source
 - » Laaja hyödyntäminen ja yhteentoimivuus välttämätöntä
- Liiketoimintamahdollisuudet lisäarvon tuottamisesta
 - Sovelluskohtaiset ontologiat
 - Semanttisen webin sovellukset



Suomalaiset semanttisen webin ontologiat (FinnONTO) –projekti



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology



- **Tavoite: kansallisen semanttisen webin infrastruktuurin kehittäminen käyntiin**
- Osahankkeet
 - Ontologioiden kehittäminen
 - Ontologiapalveluiden kehittäminen
 - Pilotointi semanttisissa portaalihankeissa
- Kaksi vaihetta
 - 1.9.2003-30.9.2005
 - 1.10.2005-31.5.2007

FinnONTO-projektin organisaatio 2003-2007



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

	Organisaatio	Jäsen
1	AAC Global Oy	Janne Nevasuo
2	AlmaMedia Oyj	Hannele Vihermaa
3	Antikvaria-ryhmä	Merja Heiskanen
4	Connexor Oy	Pasi Tapanainen
5	Elisa Oyj	Aimo Maanavilja
6	Espoon kaupunginmuseo	Liisa Ropponen
7	Geologian tutkimuskeskus GTK	Olli Rantala
8	Grip Studios Oy	Aki Kivelä
9	Helsingin kaupunginkirjasto	Erkki Lounasvuori
10	Kansalliskirjasto	Juha Hakala
11	Kansanterveyslaitos	Eija Hukka
12	Kustannus Oy Duodecim	Ilkka Kunnamo
13	Maa- ja metsätalousministeriö	Antti Vertanen
14	Maanmittauslaitos	Reino Ruotsalainen
15	M-Cult ry	Minna Tarkka
16	Museovirasto	Sirkka Valanto
17	Opetushallitus	Liisa Lind
18	Opintoluotsi-projekti/OPM	Minna Pesio-Kokko
19	Sanastokeskus TSK ry	Lena Jolkkonen
20	Siminen Meteoriitti Oy	Aki Antman
21	Sosiaaliportti-projekti/Stakes	Ani Kajander
22	Stakes (luokittelukeskus)	Matti Ojala
23	Suomalaisen kirjallisuuden seura	Lauri Harvilahti
24	Suomen maatalousmuseosäätiö	Reijo Bäckman
25	Suomen valokuvataiteen museo	Asko Mäkelä
26	TietoEnator Oyj	Riitta Alkula
27	Valtion taidemuseo	Elina Heikka
28	Valtiovarainministeriö	Riku Moisio
29	Yleisradio Oy	Tomi Nummi
1	Helsingin yliopisto (TKTL)	Eero Hyvönen
2	Helsingin yliopisto (YKL)	Lauri Carlson
3	Tampereen yliopisto	Kalervo Järvelin
4	Teles	Marko Heikkinen
5	TKK (kartografia & geoinformatiikka)	Kirsi Virrantaus
6	TKK (viestintäteknikka)	Eero Hyvönen



- 1. Julkisten **ydinontologioiden kehittäminen** open source
 - Yleinen suomalainen ontologia YSO, perustana YSA
 - YSO:aan liittyvät alakohtaiset ontologiat
- 2. Kansallisten **ontologiapalveluiden kehittäminen**
 - Toisistaan riippuvien ontologioiden rakentamisen, versioinnin ja julkaisemisen kehitysympäristö ONKI
 - Tiedon indeksointi web-palvelun avulla
 - Tiedon käsiteperustainen haku web-palvelun avulla
- 3. **Pilotointi portaalihankkeissa**
 - Useita sovellushankkeita

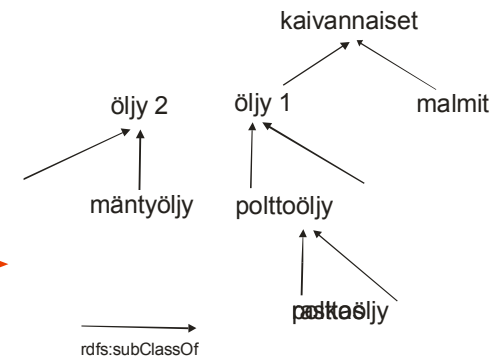
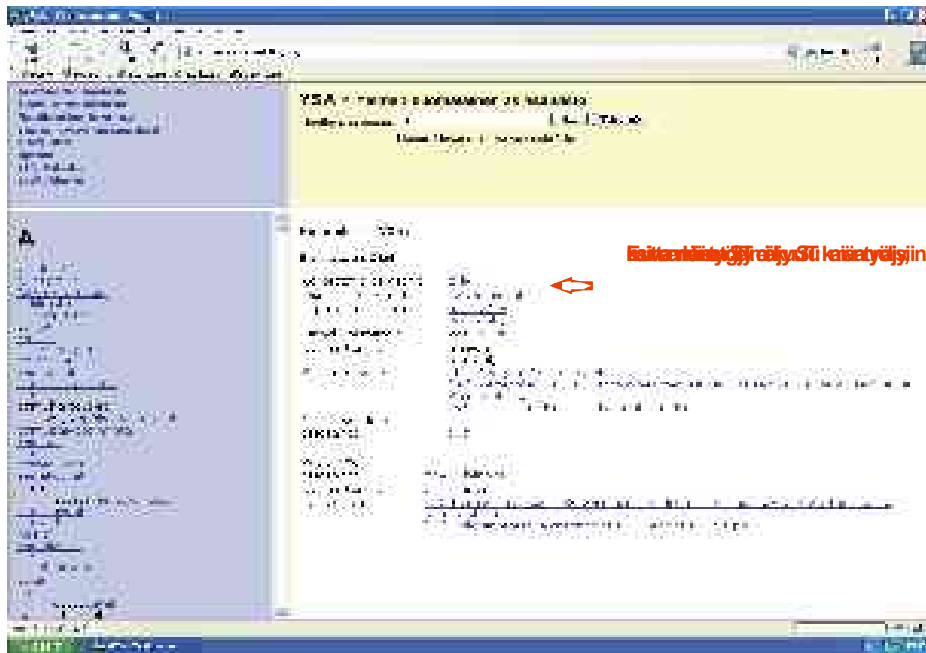
1. Julkisten ontologioiden kehittäminen



- Motto: Asiasanatesauruksista ontologioihin!
- Yleinen suomalainen ontologia YSO
 - Perustuu Yleiseen suomalaiseen asiasanastoon YSA

YSA

YSO



YSOon liittyviä muita ontologioita



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

Ontologia	Kuvaus
YSO	Yleinen suomalainen ontologia. Kansalliskirjaston Yleisen suomalaisen asiasanaston (YSA) ontologisoitu versio.
MAO	Museoviraston Museoalan asiasanastosta kehitetty Museoalan ontologia. Ensimmäinen versio käytössä MuseonSuomessa.
PAIKKO	Suomalainen paikkaontologia.
SAPO	Suomen ajallinen paikkaontologia eri aikakausina. Ensimmäisen versio n tavoitteena on kuvata Suomen kunnat ja niiden muutokset 1900-luvusta lähtien.
ICONCLASS	Kansainvälisesti käytetty, Hollannissa kehitetty taitteen kuvausjärjestelmä, jota Valtion taidemuseo on soomentanut taidemuseoiden käyttöön. Järjestelmä muunnetaan ontologiaksi ja yhdistetään YSO:aan.
VALO	Valokuvataiteen museon kolmesta asiasanastosta muodostettu ontologinen kokonaisuus.
TOIMO	Henkilöiden, yritysten, organisaatioiden yhteis toimijoiden ontologia. Ensimmäinen versio käytössä MuseoSuomessa.
TAO	Tapahtumien ja prosessien ontologia. Perustuu YSO:aan.
MATO	Matalomiesalan ontologia. Perustuu asiantun Viikin tiedokirjaston Argiforest sanastoon.



UNIVERSITY OF HELSINKI



SeCo
SOLUTIONS

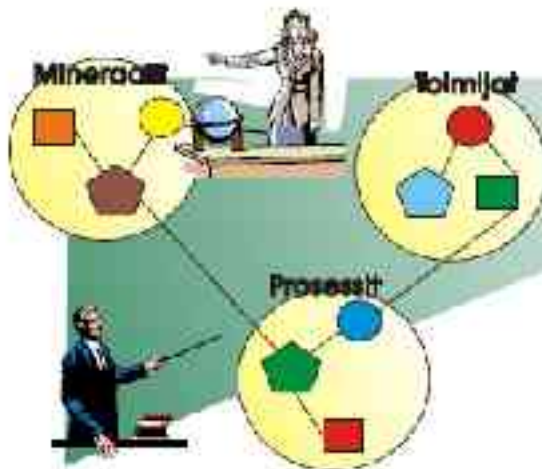
2. Kansalliset ontologiakirjastopalvelut (ONKI)



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

Ontologian kehittäjät

- Asiakas-palvelimalli tukee hallittua hajautusta
- Käsitteiden turvallinen lisääntyminen antaa mahdollisuuden jakaa kehitys halutuille tahoille
- Muutostenhallinta pitää ajantasalla muiden ontologioiden muutoksista
- Versiohistoria tarjoaa näkymän mallien historiaan ja kunkin vaiheen muutoksiin



Tiedon hakija

- Selainnäköymä käsitteisiin ja niiden välisiin suhteisiin
- Käsitteisku

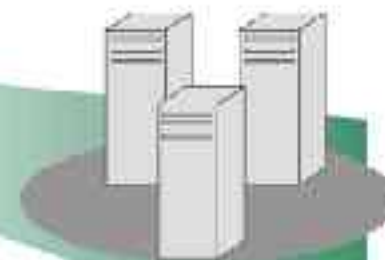


Nokia?
Palkko vai yritys?



Tiedon indeksoija

- Selain- ja Web Service-rajapinnat käsitteiden selailu-, haku- ja käyttöönottotoimintoihin, jolla automatisoinnista voidaan saada maksimaalinen hyöty



3. Pilottisovellukset (SeCo-ryhmä)



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- MuseoSuomi
 - Suomen museot semanttisessa webissä
- KulttuuriSampo
 - Suomen kulttuuri semanttisessa webissä
- Orava
 - YLE:n Klaffi-portaalin semanttinen versio
- Semanttinen Suomi.fi
- Uusia portaalihankkeita liittynyt projektiin 2005
 - Tiedonportti.fi
 - Kansalaisen terveyden edistämisen portaalii
 - Opintoluotsi.fi
 - Sosiaaliportti.fi



MuseoSuomi -- Suomen museot semanttisessa webissä



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Ideat
 - Globaali web-näkymä kokoelmatietokantoihin
 - Semanttinen hakukone + samoilu
 - Museoiden yhteinen julkaisukanava
- <http://www.museosuomi.fi>
- Kansainvälisesti palkittu sovellus
 - Semantic Web Challenge Award 2004
 - Nordic Digital Excellence in Museums 2004
 - Pääministerin kunniamaininta



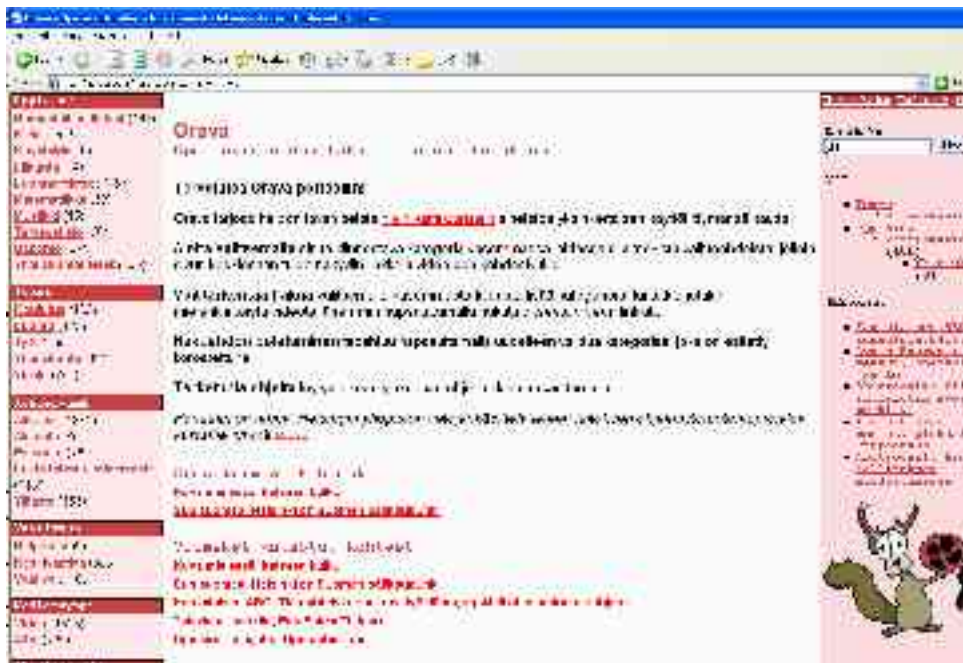
Orava: YLE:n Klaffi-portaalin semanttinen versio



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

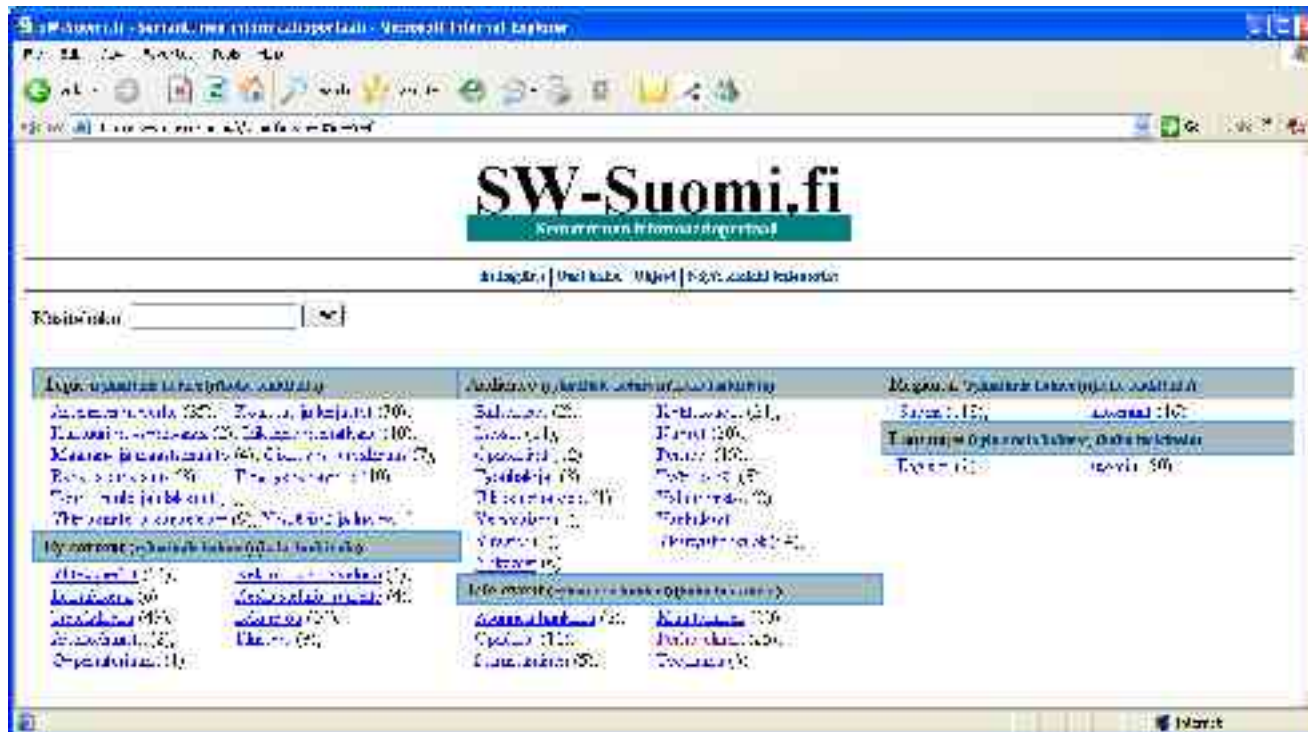
- Semanttinen oppimateriaalin haku ja selailu
 - » 2200 videota ym., Learning Object Metadata (LOM)
- Semantic autocompletion –ominaisuuden kehittäminen
- Portaalien sisältöjen yhdistäminen
 - » Linkitys MuseoSuomi-aineistoihin

<http://www.museosuomi.fi/orava>

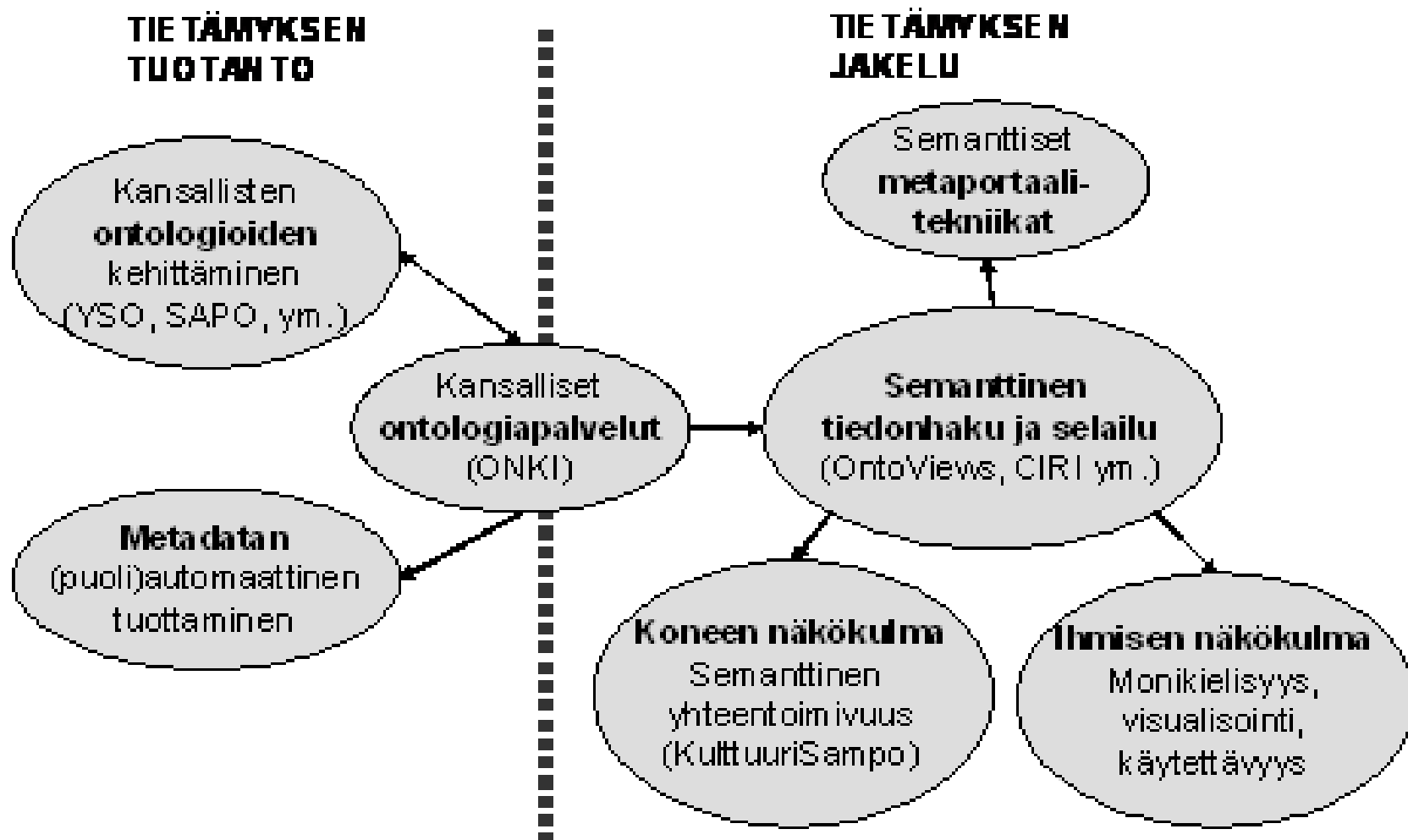




- Uusien näkökulmien (5) luominen linkkikirjastoon
 - Semanttinen moninäköymähaku ja selailu
- Eri hallinnonalojen palveluiden yhdistely ongelmakeskeisesti
 - Esim. ”Lapsen syntymä”



Tutkimuskohteita



Tekes Fenix-ohjelma: Tietämyksen hallinnan ja semanttisen webin aihieryhmä



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Rakentunut FinnONTO hankkeen ympärille
- Toimintaa
 - Älyä verkkoon – Web Intelligence
 - » Symposio Heurekassa 2004
 - 16.11.2005 Kohti suomalaista semanttista webiä
 - » Seminaari TKK:ssa
 - Digital Content Across Cultures
 - » Seminaari Pariisin Louvressa, 4-5.5.2005 (?)
 - Sähköpostilista
 - » Kaikki tervetulleita mukaan!
- Kaikki toimintaideat tervetulleita!
- Yhteydenotot
 - eero.hyvonen@tkk.fi



- Semanttinen viestintäteknikka ja web tulevat
 - » Semanttinen web = webin seuraava sukupolvi
- Kehitys edellyttää kansallista panostusta ja yhteistyötä
 - » Semanttisen webin suomalaiset ontologiat
 - » Ontologiapalvelut vapaasti käytettäviksi
 - » Pilotointi ja käytännön sovellukset
- Tärkeä ala jatkossa monessa mielessä
 - » Teollinen intressi
 - » Tekninen mahdollisuus
 - » Tieteellinen haaste
 - » Kansallinen intressi