



Ontologioiden yhdistäminen YSO:oon

Katri Seppälä, Päivi Lipsanen ja Reetta
Sinkkilä

Semanttisen laskennan tutkimusryhmä (SeCo)
Teknillinen korkeakoulu, mediatekniikan laitos;
Helsingin yliopisto, tietojenkäsittelytieteen laitos

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

Erikoisalojen asiasanastoista yhdistelmäontologioihin



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- Yleisellä suomalaisella asiasanastolla (YSA) ja erikoisalojen asiasanastoilla on huomattava määrä yhteisiä käsitteitä ja erikoisalojen asiasanastojen rinnalla käytetään usein YSA:a
 - käsitteiden väliset suhteet eivät kuitenkaan aina ole yksiselitteisiä
- yhdistelmäontologiat, joissa YSO:a on käytetty rakenteellisena mallina ja linkkinä erikoisalojen käsitteistöön
 - selkeyttävät eri ontologioiden yhteiset käsitteet
 - tuovat esiin monimerkityksiset termit ja niiden taustalla olevat eri käsitteet
 - helpottavat kokonaisuuden käyttöä ja kehitystä
 - » auttavat välttämään turhaa päällekkäisyyttä ja mahdollistavat entistä paremmin omaan käsitteistöön keskittymisen

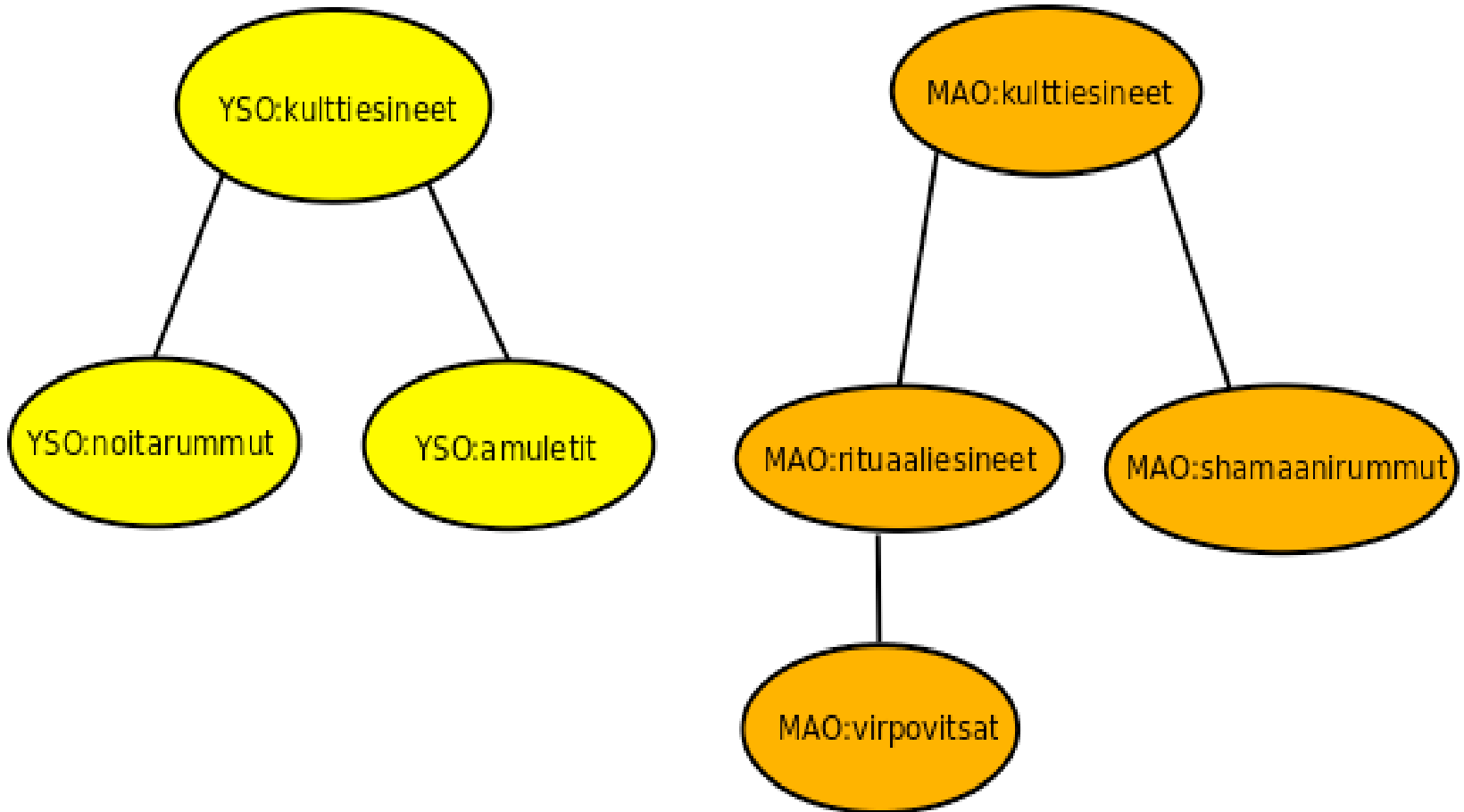


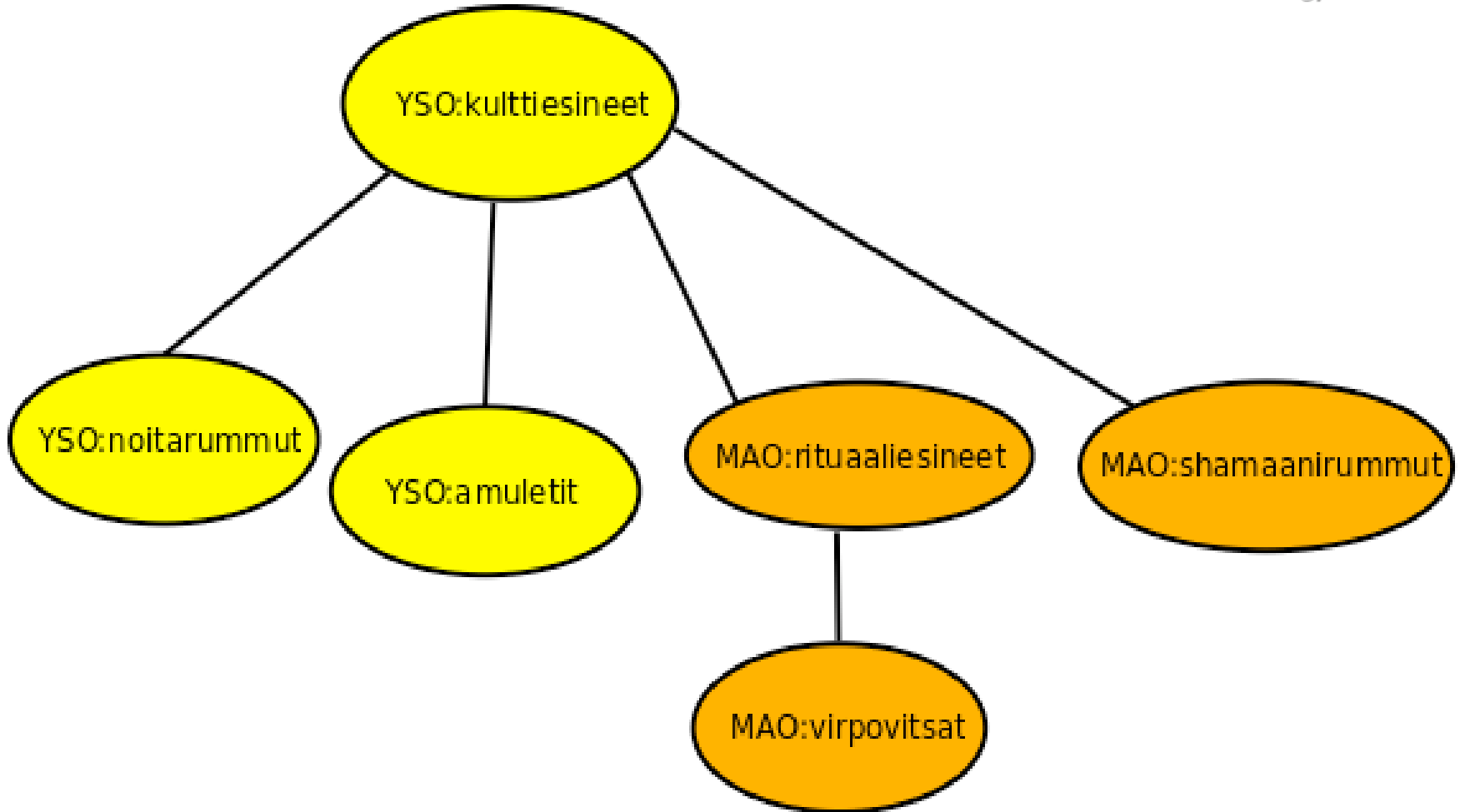
- (asia)sanojen (label) merkkijonopohjainen vertailu
 - kerätään kaikkien ontologioiden asiasanat ja muutetaan ne perusmuotoon
 - etsitään eri ontologioiden yhteiset asiasanat ja merkitään käsitteet yhteneviksi (owl:equivalent) yhteisten asiasanojen perusteella
 - » yhtenevien käsitteiden etsinnässä on mahdollista käyttää ensisijaisten asiasanojen lisäksi myös korvattuja asiasanoja/synonyymeja (altLabel) tai erikielisiä asiasanoja
 - yhdistelyssä käsitellään aina vain kahta ontologiaa kerrallaan, vaikka lopullisena tavoitteena olisi useamman ontologian yhdistäminen



- erikoisontologian ripustus YSO:oon
 - tavoitteena täydentää YSO:a erikoisontologian käsittein
 - ripustus perustuu owl:equivalent-määritteisiin
 - erikoisontologian YSO:n kanssa yhtenevien käsitteiden alakäsitteet liitetään YSO:n vastaavan käsitteen alle hierarkkiseksi alakäsitteiksi (rdfs:subClassOf)

Yhdistelmäontologiat – koneellinen yhdistely





Yhdistelmäontologiat – koneellisen yhdistelyn tarkistus ja täydennys



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

- pitääkö ehdotettu ekvivalenssi paikkansa?
 - onko sanaa käytetty eri ontologioissa samassa merkityksessä? esim. kaavat, esteettömyys
- onko aineistosta jäänyt puuttumaan ekvivalenssimerkintöjä
 - käytetäänkö eri ontologioissa eri synonyymeja? esim. applikaatiotyöt/applikointityöt, kitsch/kitsi
- mihin ei-ekvivalentit erikoisalan ontologian käsitteet sijoittuvat YSO:n hierarkiassa?
- onko mahdollista tarkentaa erikoisalan ontologian sisäistä hierarkiaa?



- Taideteollisuusalan ontologia (TAO)
 - perustuu Muotoilun ja viestinnän asiasanastoon
 - n. 3 000 käsitettä suomen kielellä
 - aihealueet: muotoilu, viestintä, käsi- ja taideteollisuus
 - asiasanaston kehittäjä: Kuopion Muotoiluakatemia, Savonia-ammattikorkeakoulu
 - ontologiatyön yhteistyökumppani: Taideteollinen korkeakoulu



- Museoalan ontologia (MAO)
 - perustuu Museoalan asiasanastoon sekä MuseoSuomi-palvelussa julkaistuu ontologiaan
 - n. 6 800 käsitettä suomen kielellä
 - aihealueet: aineellinen kansatiede, arkeologia, konservointi, kulttuuri-, meri- ja rakennushistoria, museologia, numismatiikka
 - asiasanaston kehittäjä ja ontologiatyön yhteistyökumppani: Museovirasto



- Valokuvausalan ontologia (VALO)
 - perustuu Valokuvan asiasanastoon ja Suomen valokuvataiteen museossa kerättyihin valokuvaustekniikan ja valokuvailmaisun käsitteisiin
 - n. 2 000 käsitettä suomen kielellä
 - asiasanaston kehittäjä ja ontologiatyön yhteistyökumppani: Suomen valokuvataiteen museo



- Agriforest-ontologia (AFO)
 - perustuu Agriforest-asiasanastoon
 - n. 7 000 käsitettä
 - » käsitelty toistaiseksi vain suomen kielen näkökulmasta
 - » sisältää myös englanninkielisiä vastineita suomenkielisille termeille
 - aihealueet: maatalous-, metsä-, puu-, elintarvike-, kotitalous-, kuluttaja- ja ympäristöala
 - asiasanaston ja ontologian kehittäjä: Viikin tiedekirjasto

YSO + AFO – puolukka = kasvi



pysyvä (yso)
fyysinen objekti (yso)
orgaaninen rakenne (yso)
eliöt (yso, afo)
kasvit (yso, afo)
kasvit taksonin mukaan (yso, afo)
putkilokasvit (yso, afo)
siemenkasvit (yso, afo)
koppisiemeniset (yso, afo)
kaksisirkkaiset (yso, afo)
kanervakasvit (yso, afo)
puolukat (afo)
puolukka_1 (afo)

YSO + AFO – puolukka = marja



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

pysyvä (yso)

fyysinen objekti (yso)

orgaaninen rakenne (yso)

kasvinosat (yso, afo)

hedelmät (yso, afo)

marjat (yso, afo)

luonnonmarjat (yso, afo)

puolukka (yso) = puolukka_2 (afo)

YSO + AFO – vektorit_1 – vektorit_2



HELSINKI UNIVERSITY OF TECHNOLOGY
Media Technology

abstrakti (yso)

vektorit_1 (yso)

vektori (yso) – suure, jolla on tietty suuruus ja suunta (Kielitoimiston sanakirja)

pysyvä (yso)

ominaisuudet (yso)

ulkoiset ominaisuudet (yso)

roolit (yso)

roolit vaikutuksen mukaan (yso)

vektorit_2 (yso) = vektorit (afo)

vektorit (afo) = väline, joka kuljettaa tietyn kappaleen paikasta toiseen,
esim. plasmidi geenitekniikassa tai loisen väli-isäntä (Biologian sanakirja)



- yhdistelmäontologioiden testiversioita on julkaistu ONKI-palvelimella
 - Taideteollisuuden ja Museoalan ontologiat + YSO
 - » <http://www.yso.fi/onki/tao/>
 - Valokuvausalan ontologia + YSO
 - » <http://www.yso.fi/onki/valo/>
 - Agriforest-ontologia + YSO
 - » julkaistaan kevään aikana
- kehitystyötä jatketaan SemWeb 2.0-hankkeessa