

# Sisällönkuvailun tulevaisuus: YSA vai YSO?

Eeva Kärki  
Kansalliskirjasto  
22.1.2009

## Agenda

- YSA
- YSO
- YSA – YSO: eroja
- Selvitettävää
- Osoitteita

## Tesaurukset ja ontologiat

- molemmat ovat käsitejärjestelmän kuvauksia
- tesaurus-suhteet/ontologia-suhteet
- asiasanastot = ihmisen tulkittavissa olevia
- ontologiat = koneen (ja ihmisen?) tulkittavissa olevia

## YSA eli Yleinen suomalainen asiasanasto (1)

- n. **27 000** asiasanaa (näistä n. **5000** maantieteellistä nimeä)
- kaikki tiedonalat
- ei henkilön- eikä yhteisönnimiä
- yleissanasto (epätasaisuutta...)
- tesaurus-rakenne
- aakkosellinen ja systemaattinen osa
- ylläpito: Kansalliskirjasto (1 htv)
- ”tekninen” ylläpito: Fennica-tietokanta (Voyager)

## YSA (2)

### **Käyttö**

- Fennica (Suomen kansallisbibliografia)
- kaikki yleiset kirjastot
- suurin osa tieteellisistä kirjastoista (YSA:n lisäksi monet tieteelliset kirjastot käyttävät oman alansa erikoissanastoja)
- arkistot
- museot (Museoalan asiasanaston lisäksi)

## YSA (3)

### **Tesaurus-rakenne (YSA)**

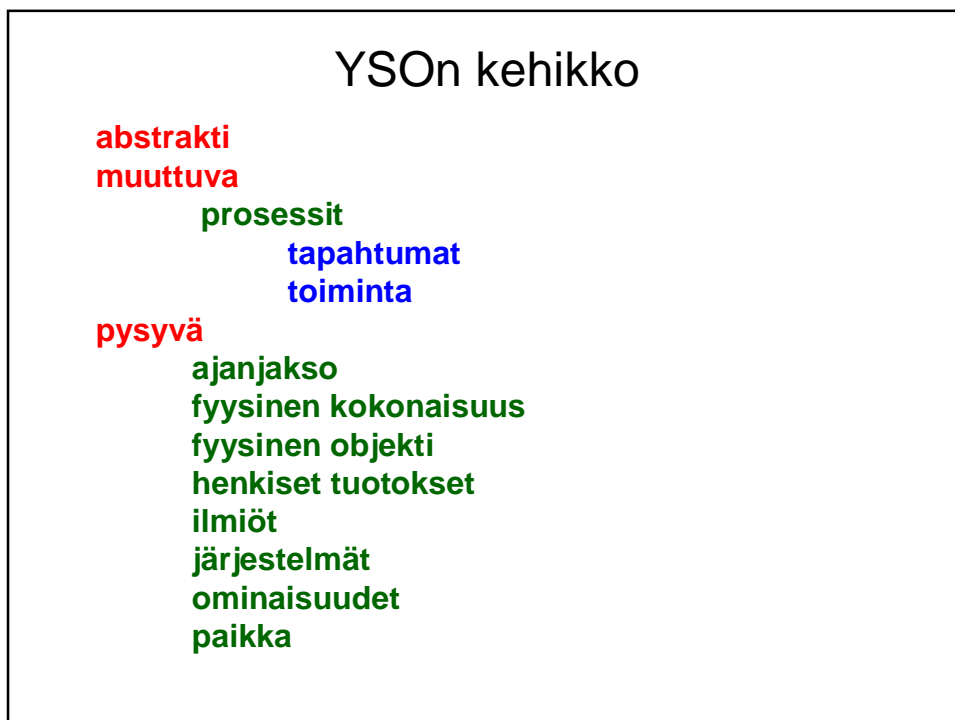
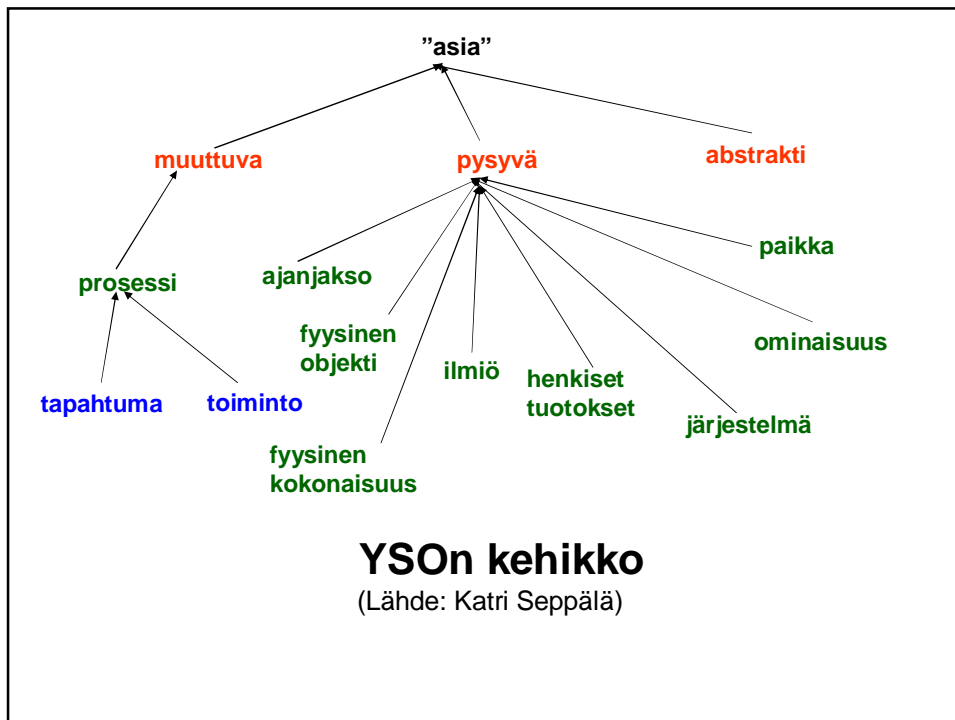
- KÄYTÄ = viittaus ohjaustermistä asiasanaan (VESA = asiasanaa xxx ei käytetä)
- KT = korvaa termit
- LT = laajemmat termit
- ST = suppeammat termit
- RT = rinnakkaistermit
- Huomautus
- Lähde
- Ryhmäkoodi ( VESA = kuuluu ryhmiin)
- Kaikki termisuhteet ovat aina vastavuoroisia

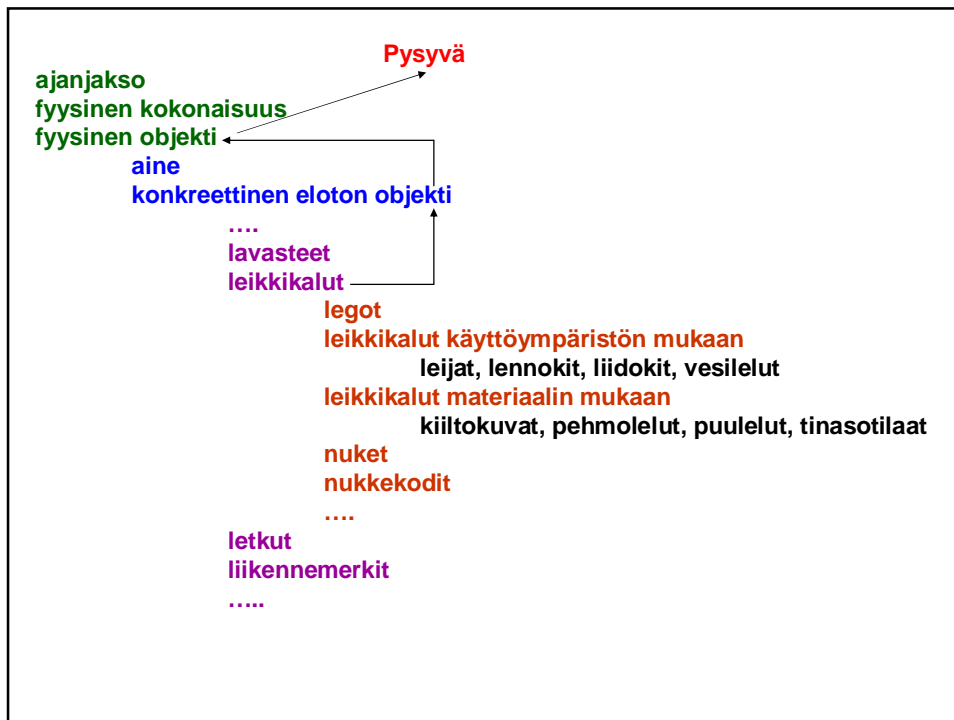
## YSO eli Yleinen suomalainen ontologia(1)

- FinnOnto-osaprojekti (Tekes) 2004-2007, jatko projekti 2008-2010
- perustuu YSAan (**puuttuu** vuoden 2006, 2007 ja 2008 YSA-termit eli **yli 2000 termiä**)
- n. 20 000 termiä
  - **ei maantieteellisiä nimiä**
- kieli: suomi
  - ruotsinkieliset vastineet (Allärs)
  - englanninkieliset vastineet
    - poimittu Helsingin kaupunginkirjaston luokituskaavion englanninkielisen version hakemistosta; vastineissa virheitä, koska termejä ei ole tarkistettu erikoisalojen sanastoista/sanakirjoista
  - ruotsin- ja englanninkielisten vastineiden ontologista rakennetta ei ole tarkistettu

## YSO (2)

- jatko projekti (2008-2010):
  - YSA:n puuttuvien käsitteiden vieminen YSOon
  - Allärsin ja Musa/Cillan ontologisoiminen
  - entä englanninkieliset termit?
- YSOa käytetään myös rakenteellisena mallina ja liittymäkohtana erikoisalojen ontologioihin, kuten
  - AFO (Agriforestista tehty ontologia)
  - VALO (Valokuva-alan ontologia)
  - MAO (Museoalan ontologia)
- osoite: <http://www.yso.fi/onki/yso/>





## YSOn kehikko

### ST-suhteet

- generiset suhteet (laji – alalaji)
  - voimalat - lämmitysvoimalat
- partitiiviset suhteet (kokonaisuus – osa)
  - kasvinosat – juuri, kukka jne.
  - esitetään eri tavalla kuin generiset ST-suhteet
  - näitä ei ole vielä tehty YSOon (eli ovat nyt tavallisia ST-suhteita)

## YSA – YSO (1)

### Sanojen väliset suhteet

- **YSAssa** asiasanan ympärille rakennetaan tesaurus-suhteet aikaisemmin YSAssa oleviin asiasanoihin
  - kaikki YSAssa esiintyvät sanat ovat ”todellisia” asiasanoja
  - monilla sanoilla ei ole hierarkkisia suhteita, vain assosiatiivisia suhteita (RT-suhde)
  - joillakin sanoilla ei ole mitään suhteita
  - käsitejärjestelmäkokonaisuudet melko pieniä

## YSA – YSO (2)

### Sanojen väliset suhteet (jatkuu)

- **YSOssa** on valmis kehikko, johon kukin sana yritetään sijoittaa mahdollisimman loogisesti
  - ontologiassa paljon hierarkian vaatimia ns. ”aputasosanoja”, jotka eivät ole asiasanoja
  - käsitekehikko monimutkainen ja tarkka
  - kaikilla sanoilla/termeillä on joku hierarkkinen suhde
  - pitkiä hierarkioita (esim. kissa)
  - monet käsitteet vaikea sijoittaa kehikkoon loogisesti (ongelma varsinkin ns. ”pehmeillä tiedonaloilla”) -> vaarana, että kokonaisuus on epäjohdonmukainen (esim. avioero – asumusero)

## YSA – YSO (3)

### Sanojen monimerkityksisyys yms.

- **YSAssa** ratkaistu kahdella tavalla (esim. kaanonit, kurkku, messut)
- **YSOssa** pyritty ottamaan esille mahdollisimman paljon monimerkityksisyyttä (kaanonit, kurkku, messut)
  - kuinka ”syvälle” polysemiaan kannattaa mennä?
- **YSOssa** osa termeistä esitetty myös eri ”funktioissa” (esim. lapset, lapset (kooste))

### Maantieteelliset nimet

- **YSAssa** mukana n. 5000 paikannimeä (myös luonnonmaantieteellisiä nimiä, koti- ja ulkomaisia)
- **YSOssa** ei maantieteellisiä nimiä, maantieteelliset nimet paikkaontologioissa (SUO, SAPO; sisältävät suomalaisia paikannimiä)

## YSA – YSO (4)

### Alanmukaiset ryhmät

- **YSAssa** voidaan selata alanmukaisia ryhmiä
- **YSOssa** ei voida selata alanmukaisia ryhmiä (ainakaan vielä)

### Vapaa indeksointi

- **YSAssa** ohjeet vapaan indeksoinnin sanaryhmien yhteydessä
- **YSOssa** asia ratkaisematta

## Selvitettävää: kysymyksiä (1)

- **Miksi** pitäisi siirtyä asiasanoista ontologioihin?
- Mitä todellista **lisäarvoa** ontologiat tuovat sisällönkuvailuun ja tiedonhakuun?
- Ontologiat luotu lähinnä koneymmärrettäväksi, soveltuvatko ne **ihmisen tulkittaviksi**?
- Ovatko ontologiat **sisällöltään** niin valmiita, että niitä voidaan ryhtyä käyttämään?
- **Maantieteelliset** nimet: tulevatko SUO ja SAPO osaksi KOKO-ontologiaa? Entä ulkomaiset paikannimet?

## Selvitettävää: kysymyksiä (2)

- Miten käytetään niitä asiasanoja, jotka **eivät ole missään ontologioissa**?
  - esim. vapaan indeksoinnin sanat, numeeriset ajanmääreet
- Koska asiasanaa ei voi käyttää muussa kuin ontologian esittämässä yhteydessä, miten menetellään, jos ko. **sanaa tarvitaankin ihan muissa yhteyksissä**?
  - onko otettava yhteys ontologian ylläpitäjään?
- Onko kaikki **käytettävät asiasanastot** ontologisoitava?
  - resurssit? Sisällöllinen ja tekninen ylläpito?
- Entä **ulkomaisten** asiasanastojen ja ontologioiden käyttö?

## Selvitettävää: kysymyksiä (3)

- Onko **tekniisiä valmiuksia** ryhtyä käyttämään ontologioita sisällönkuvailussa ja tiedonhaussa (eri kirjastojärjestelmät )?
  - termien poiminta ontologiasta pitää olla niin **vaivatonta**, ettei se hidasta sisällönkuvailutyötä
  - miten käy ketjutusten?
- Jos ontologioita ryhdytään käyttämään asiasanoituksessa, pystytäänkö aikaisemmat indeksoinnissa **käytetyt asiasanat** kytkemään ontologian yhteyteen?
  - jos ei, mitä se vaikuttaa tiedonhakuun?

## Osoitteita

- <http://vesa.kansalliskirjasto.fi/> (pääsee sekä uuteen että vanhaan VESAn käyttöliittymään)
- <http://www.yso.fi/onki/yso/> (YSO)
- <http://demo.seco.tkk.fi/onkipaikka/> (SUO = Suomalainen paikkaontologia)
- <http://www.yso.fi/onki/sapo/?l=fi> (SAPO = Suomen ajallinen paikkaontologia)
- <http://www.yso.fi/onki/koko/?l=fi> (KOKO-ontologia)
- <http://www.yso.fi/?l=fi> (Ontologiapalvelin, linkit FinnONTO-projektissa työstettyihin ontologioihin)
- <http://www.yso.fi/lusto-demo.html> (Kantapuu-museotietokanta, hakudemo)
- <http://wordnet.princeton.edu/> (WordNet)