



ONKI-projekti

KERTTU

Taustaa

- **FinnONTO -- Suomalaiset semanttisen webin ontologiat (2003-2012)**
 - Aalto-yliopiston semanttisen laskennan ryhmän projekti, jossa luotiin pohjaa suomalaiselle semanttisen webin / yhdistetyn tiedon (Linked Data) infrastruktuurille
 - Tekesin rahoittamaan hankkeeseen osallistui kaiken kaikkiaan n. 50 yritystä ja organisaatiota
- **Ontologiatyön ja siihen liittyvien palvelujen jatkaminen koettu erittäin tärkeäksi**

ONKI-projekti

- **ONKI-projekti on Kansalliskirjaston, Opetus- ja kulttuuriministeriön ja Valtiovarainministeriön yhteisprojekti**
- **Ensimmäinen osuus päättyy 31.3.2014**
- **Suunniteltu jatkumaan vuoteen 2017**

Tekijät

- **Projektipäällikkö Matias Frosterus**
- **Tietojärjestelmäasiantuntijat Sini Pessala, Osma Suominen ja Henri Ylikotila**
- **Tietoasiantuntijat Mikko Lappalainen, Susanna Nykyri ja Tuomas Palonen**
- **Asiasanastoasiantuntijat Mirja Anttila ja Liisa Virtanen**
- **Suunnittelupäällikkö Annu Jauhiainen**

Ontologiat

- **Ontologia on jonkin aihealueen käsitteiden eksplisiittinen määrittely**
 - Käsitehierarkia, joka kuvaa käsitteiden väliset suhteet

Ontologiat

- **Hiljainen tieto**
- **Yleistieto**

- **Koneymmärrettävä muoto**

Ontologiat

- Jos eri sovellusalat ja niillä toimivat organisaatiot käyttävät tiedon indeksoinnissa **samoja käsitetunnisteita (URI) samoista asioista**, niin järjestelmien integrointi helpottuu.

Ontologiat ja asiasanastot

- **Termien sijasta käsitteitä**
 - Kieliriippumattomia
 - Pakottaa eksplikoimaan
 - Peilaukset käsitteiden välillä selkeämpiä
 - Etenkin monikieliset peilaukset

Ontologiat ja asiasanastot

- **Eksaktit suhteet**

- Ominaisuuksien periytyminen

- Tyypitys

- Esimerkiksi, paikkatietokenttään halutaan vain paikkoja, jolloin Nokiaa yrityksenä ei voi vahingossa syöttää siihen

- **Koneluettavuus**

- Konepäättely

- Esimerkiksi kyselynlaajennus: "työkalut"-haku löytää myös sahat ja vasarat

Tavoitteet

- **Tarjota pysyvä, keskitetty, Kansalliskirjaston ylläpitämä ontologiapalvelu**

Tavoitteet

- **Tarjota pysyvä, keskitetty, Kansalliskirjaston ylläpitämä ontologiapalvelu**
 - Ontologioiden julkaisu
 - Ontologioiden käyttö
 - Selailu
 - Rajapinnat
 - Ohjeita ja tietoa
 - Perustuu avoimeen lähdekoodiin

ONKI-palvelu:

Julkaisuputki

Palvelimet

Toimintata-
poja, tukea

ONKI-
Light

YSO
ALLSO

ONKI-palvelu:

Julkaisuputki

Palvelimet

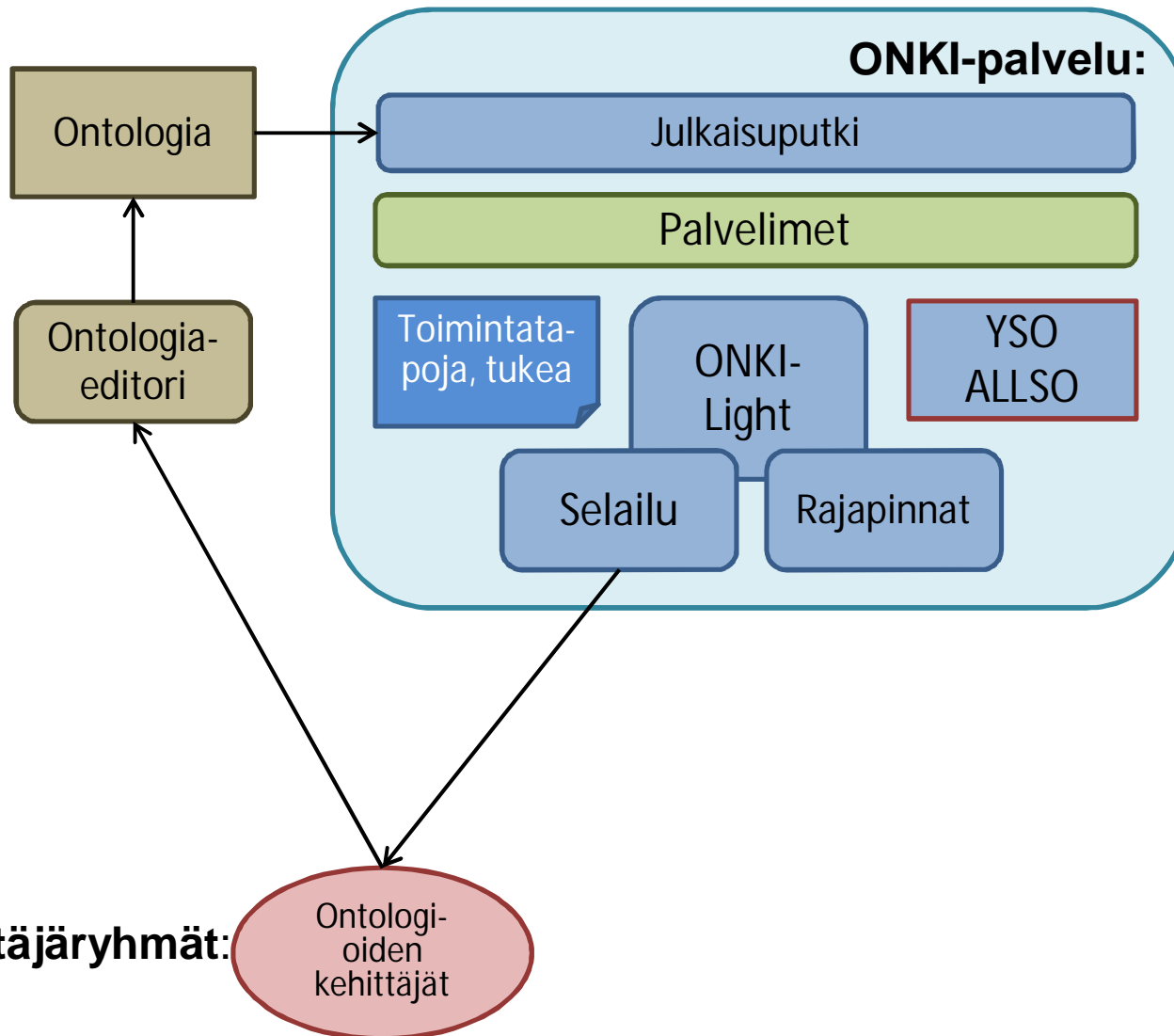
Toimintata-
poja, tukea

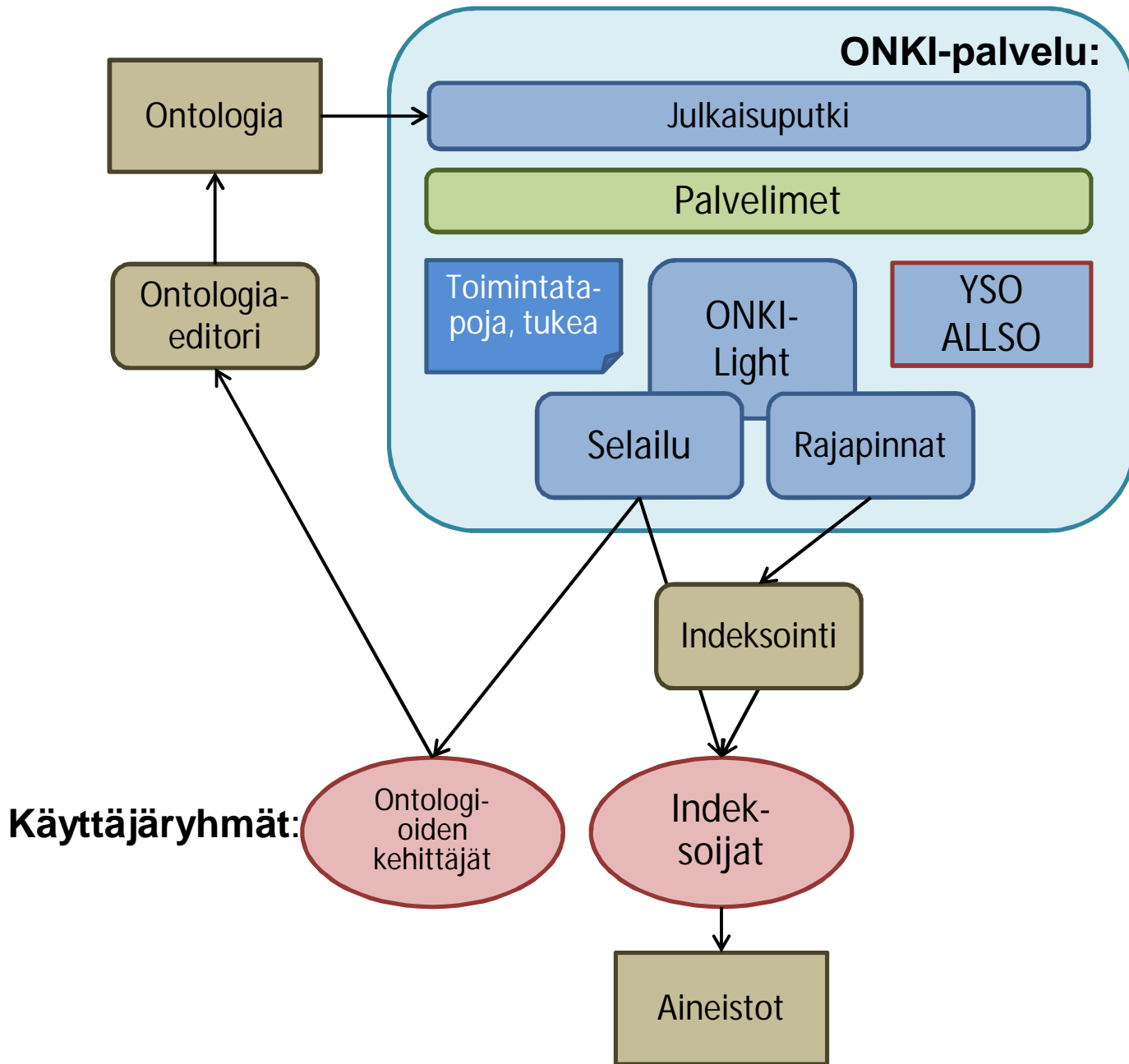
ONKI-
Light

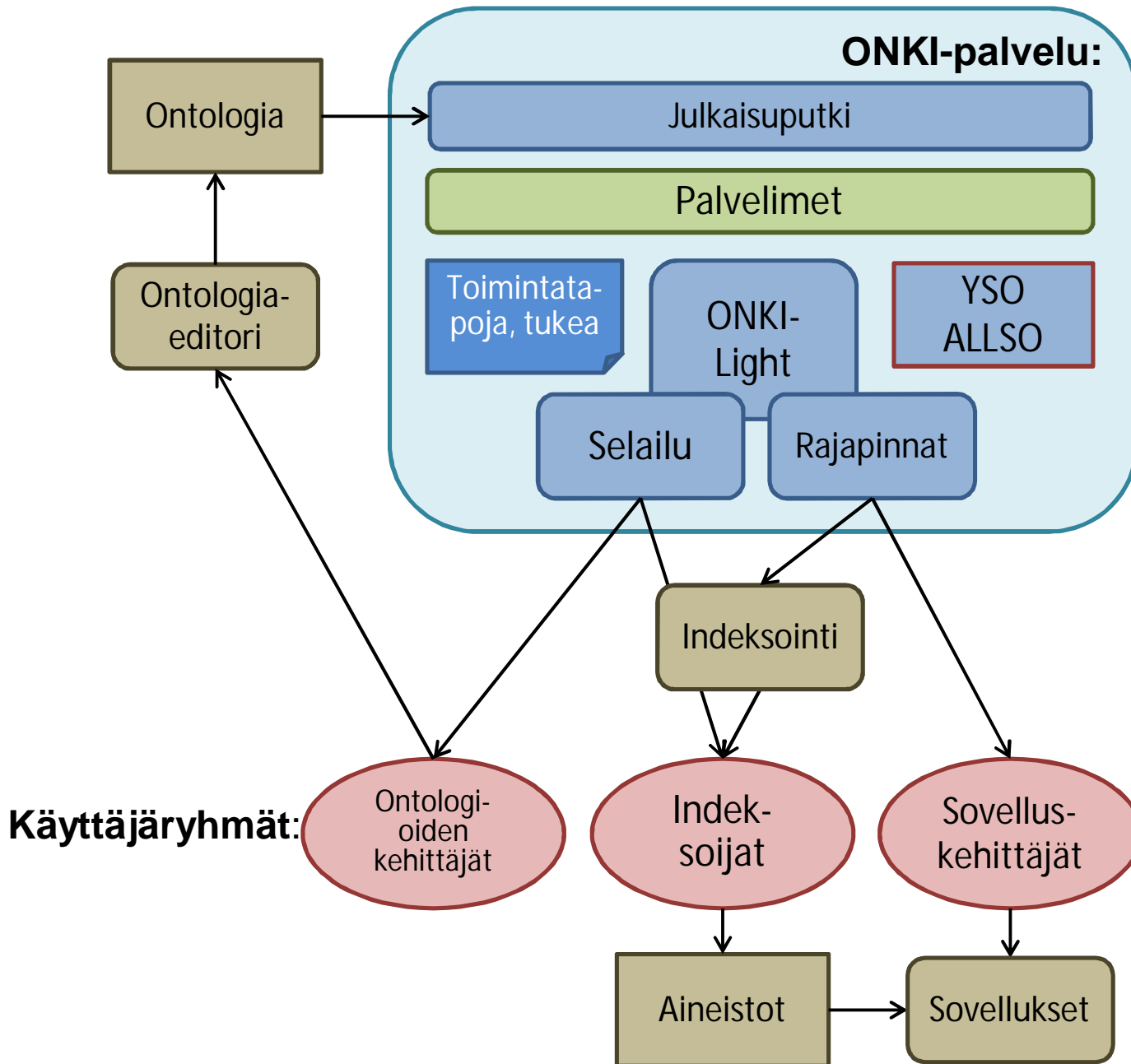
YSO
ALLSO

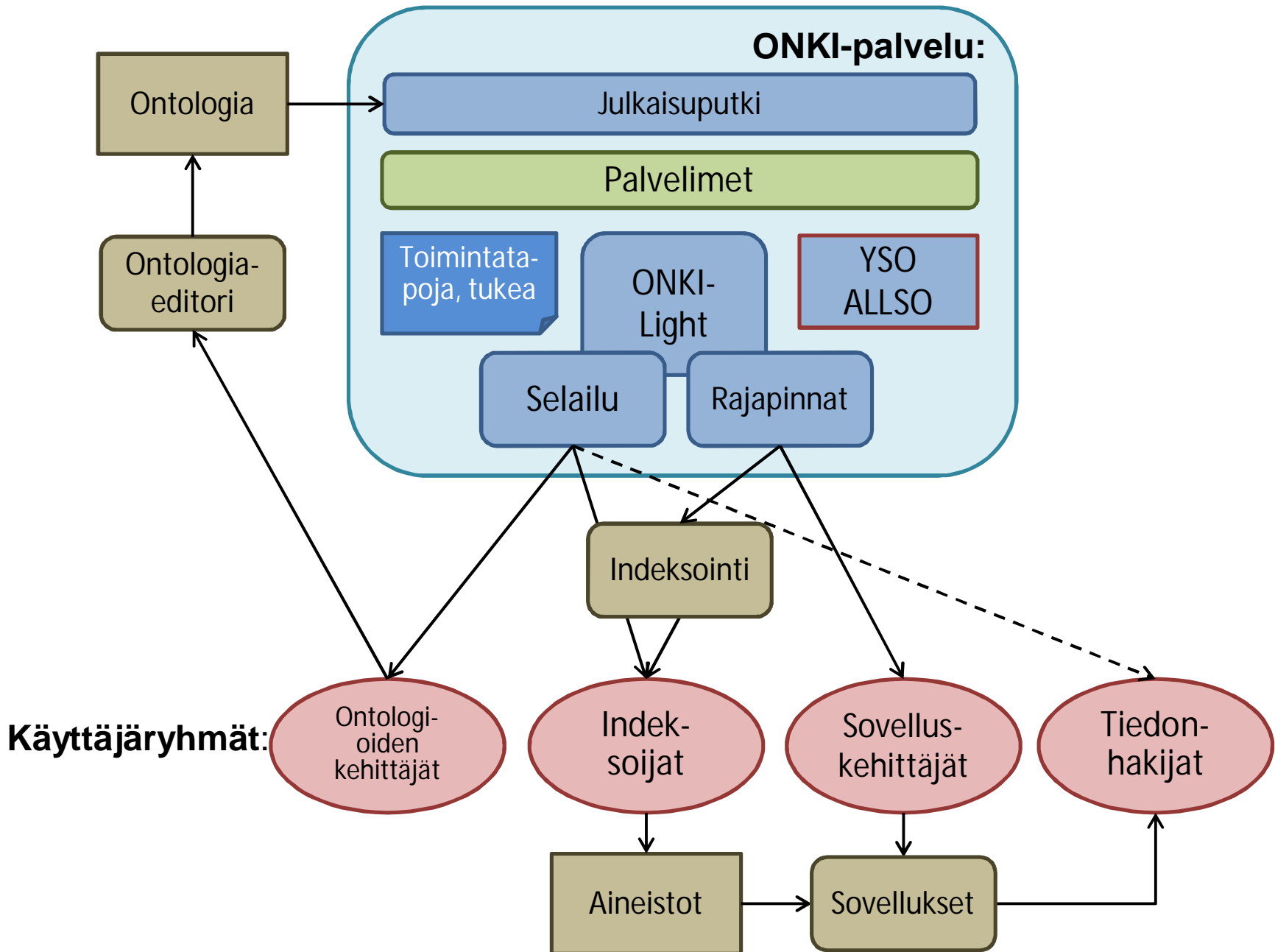
Selailu

Rajapinnat

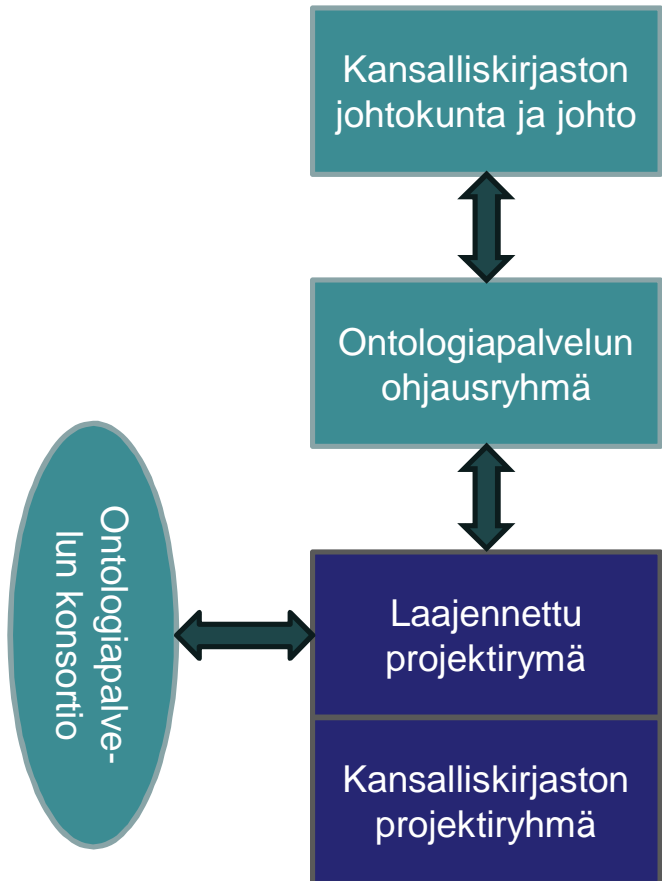








Projektin organisaatorakenteesta



- **Ohjausryhmä**
 - Linjaa kehitystä, käsittelee laajakantoisia ja periaatteellisia päätöksiä, hyväksyy toimintasuunnitelman
- **Laajennettu projektiryhmä**
 - Työstää käytännön ratkaisuja ja kehitystä
 - Edistää tietohallinnon kuvailutyön yhtenäistämistä
- **ONKI-konsortio**
 - ONKIn käyttäjistä ja sisällöntuottajista koostuva ryhmittymä, joka koordinoi ontologioiden käyttöä ja kehitystä

YSO

- **Yleinen Suomalainen Ontologia**
 - YSAn ontologisoitu versio
 - Hierarkiaa täydennetty
 - Asiasanastotermeistä ontologiseksi käsitteiksi
 - Maantieteelliset käsitteet poistettu
 - Toimii yleisontologiana
 - Tarjoaa peruskäsitteistön

KOKO

- **Yhdistelmä YSOsta ja erikoisalojen ontologioista**
 - Erikoisalojen ontologioiden ei tarvitse suoraan linkittyä toisiinsa, vaan riittää, että ne kaikki linkitetään YSOon
 - Kehittäjälle KOKO on joukko ontologioita
 - Käyttäjälle KOKO esittäytyy yhtenä ontologiana

Lisätietoja

- onki-posti@helsinki.fi
- **Julkinen wiki:**
<https://wiki.helsinki.fi/display/ONKI/>

PROJEKTISUUNNITELMA

Tutkimushankkeesta tuotantoprojektiksi

- **Vanhojan ratkaisujen kriittinen tarkastelu**
 - Osa päätöksistä on tehty kauan sitten
 - Fokus on muuttunut
 - Tutkimisesta toimimiseen

Tutkimushankkeesta tuotantoprojektiksi

- **ONKI**
 - Vakaus, nopeus, käytettävyys
 - Käyttötarpeiden selvitys ja niihin vastaaminen

Tutkimushankkeesta tuotantoprojektiksi

- **YSO**
 - Käyttötarpeiden selvittäminen
 - Kriittinen tarkastelu vastaako niihin
 - Mahdolliset hierarkiamuutokset
 - Mahdolliset rakennemuutokset (skeema)
 - Jatkokehitys

Projektin ala ja sen rajoitukset

- **Projekti päättyy 31.3.2014**
- **Projektin päämääränä on vakaa tilanne ja kehitystarpeiden tunnistaminen ja niiden toteutuksen suunnittelu**
- **Projekti tarjoaa asiakkaille palvelutasoltaan nykyistä vastaavan ONKI-ontologiakirjastopalvelun**
- **Projekti kehittää ontologioiden ylläpitäjien työvälineitä ja prosesseja**
- **Projektin aikana palveluun ei oteta mukaan uusia ontologioita**
- **Projektin pääasiallinen kohderyhmä on julkinen sektori, etenkin ne organisaatiot, jotka jo hyödyntävät ONKI-palvelua**
- **Projektin aikana ontologiapalvelu ei tarjoa erityispalveluita ontologioiden tutkimukseen**

Projektin hallinnolliset tehtävät

- **Projektin ohjausjärjestelmän perustaminen**
- **projektin työntekijöiden rekrytointi**
- ontologiapalvelukonsortion perustaminen
- projektista viestiminen ja konsortion markkinointi
- **projektihallinnon perustaminen**
- **tarkennetun budjetin laatiminen**
- **Tarkennettu projektisuunnitelma vuodelle 2013**
- **Tarkennettu projektisuunnitelma vuosille 2013 - 2017**
- **ONKI-palvelun yhteistoiminta- ja rahoitusmalli**
- **Tiedottaminen projektista**

Projektin ONKI-palvelun tekniseen toteuttamiseen liittyvät tehtävät

- ontologiapalvelun tavoitetilan ja palvelukuvauksen laatiminen
- ontologiapalvelun laitteistoympäristön perustaminen
- ontologiapalvelun olemassa olevien osien pystyttäminen Kansalliskirjastoon
- ontologiapalvelun arkkitehtuurin ja sovellusten nykytilan kuvaus ja arviointi
- ontologiapalvelun teknisen kehittämisen suunnitelma
- ONKI-palvelimen ylläpito vastuun siirryttyä Kansalliskirjastoon
- Uusi, tuotantokäyttöön tarkoitettu palvelinympäristö Kansalliskirjastossa ja ONKI-palvelu toiminnassa tässä ympäristössä
- ontologioiden ylläpidon työvälineiden ja rajapintojen kehittämissuunnitelma sekä sen mukaisten toimenpiteiden käynnistäminen ja ratkaisujen pilotointi.

Projektin sisällölliset tehtävät

- YSO-ontologian nykyisen rakenteen arviointi ja kehittämissuunnitelman laatiminen
- YSO-ontologian rakenteen uudistaminen kehittämissuunnitelman mukaisesti
- YSO:n jatkuva ylläpito Yleisen suomalaisen asiasanaston (YSA) ylläpidon kanssa
- **ALLSO:n (Allmän svensk ontologi) ylläpidon hankinta Åbo Akademiilta**
- YSO:n käsitteiden peilaaminen johonkin kansainväliseen ontologiaan ja englanninkielisen käännöksen uudistus
- ontologioiden lisenssien yhtenäistämisen suunnittelu
- suunnitelma ontologioiden julkaisemiseksi avoimena linkitettyinä datana
- YSO ja erikoisontologiat on yhteismitallistettu skeemojen osalta
- JUPOn ja JUHOn ontologiamuotoon liittyvä ylläpito, tuki ja koulutus otettu Kansalliskirjaston toiminnaksi

Tuotokset ja tärkeimmät välitapit

- **1/2013-7/2013**
 - Tarkennettu projektisuunnitelma vuodelle 2013
 - Uusi, tuotantokäyttöön tarkoitettu palvelinympäristö Kansalliskirjastossa ja Onki testikäytössä tässä ympäristössä
- **8/2013-11/2013**
 - Ontologioiden ylläpidon työvälineiden ja rajapintojen kehittämissuunnitelma
 - Ontologioiden ja ontologiapalvelun käyttäjätestauksen raportit
- **12/2013**
 - Tarkennettu projektisuunnitelma vuosille 2014-2017
 - Onki-palvelun yhteistoiminta- ja rahoitusmalli
 - onki.fi siirtyä viittaamaan Kansalliskirjaston palvelimille
- **1/2014-3/2014**
 - Julkishallinnon palveluiden ontologian JUPOn ja Julkishallinnon ontologian JUHOn ylläpidon, tuen ja koulutuksen konsepti
 - Konsortion toimintaperiaatteet laadittu ja kommenttikierros käynnistetty
 - YSO ja erikoisontologiat yhteismitallistettu rakenteen osalta

Työpaketit

- **ONKI TP 1 Ontologiaeditori**
 - Protegesta eroon
 - Vaihtoehtojen kartoitus
 - Raportti vertailujen tuloksista kaikkien käyttöön
 - Raportti on jo valmis ja julkaistu
- **ONKI TP 2 ONKI-palvelun pystytys**
 - Ensimmäinen versio jo toiminnassa Kansalliskirjaston palvelinympäristössä
 - kansalliskirjasto.onki.fi -> onki.fi
 - Vuoden loppuun mennessä

Työpaketit

- **ONKI TP 3 ONKIn testaus**
 - Käyttäjättestaus annotoijilla
 - Syksyn loppuun mennessä
- **ONKI TP 4 ONKI-palvelun ohjelmistotestaus**
 - Koneellinen testaus-järjestelmä uusien versioiden tekniseen toiminnan varmistamiseen
 - Kesän loppuun mennessä

Työpaketit

- **ONKI TP 5 YSON nykytilan arviointi**
 - Dokumentointi
 - Vertailu kansainvälisiin ratkaisuihin
 - Muutostarpeen selvitys ja suunnitelma toteuttamisesta
 - Muutosten toteuttamista
 - Vuoden loppuun mennessä
- **ONKI TP 6 Jatkosuunnitelma**
 - Alustava versio syksyllä
 - Lopullinen versio vuoden loppuun mennessä

Työpaketit

- **ONKI TP 7 Projektin hallinto**
 - Ohjausryhmän ja laajennetun projektiryhmän muodostaminen
 - Ensimmäiset kokoukset elokuussa
 - Laajennetun projektiryhmän ensimmäinen kokous siirtyi syyskuulle
 - Konsortion perustava kokous marraskuussa
- **ONKI TP 8 Viestintä**

TILANNEKATSAUKSET

kansalliskirjasto.onki.fi

- **Yleisiä parannuksia**
 - Nopeutettu toimintaa tapauksissa, joissa yhdellä käsitteellä on erittäin paljon alakäsitteitä
 - Palvelinympäristö on myös itsessään nopeampi
 - Korjattu erilaisia virheitä
 - Lisätty selkeyttäviä käyttöliittymäelementtejä
 - Palautteen perusteella
 - Mm. värejä, ohjeita, ominaisuuksia

kansalliskirjasto.onki.fi

- **Koneellinen testaus**
 - Tarkoituksena vähentää manuaalisen testauksen tarvetta aina kun uusi versio ONKI Lightista tuotetaan
 - Parantaa vakautta ja vähentää virheellisten versioiden mahdollisuutta
 - Selenium-testiympäristössä Javalla
 - TestNG-testikehyksessä
 - Testit ajetaan paikallisesti
 - Automatisointi myöhemmin
 - Kuormitustestit tekeillä

kansalliskirjasto.onki.fi

- **Seuraavaksi**
 - Vaihtoehtoinen hierarkianäkymä
 - Tapa hahmottaa ja selailla ontologiaa kokonaisuutena
 - Koodin selkeyttäminen
 - ”tutkimus”-tyyppisestä koodista tuotantotyyppiseen
 - Rajapintojen laajentaminen
 - Valikoitujen ONKI3:n rajapintojen toteuttaminen
 - Reagointi käyttäjätesteihin

Ontologiaeditorien vertailu

Tuomas Palonen

Ontologioiden vertailu

- **Ontologiaeditorilla ylläpidetään ontologioita: lisätään ja muokataan käsitteitä.**
- **Editorit jakautuivat yhtäältä ilmaisiin ja kaupallisiin, toisaalta paikallisesti asennettaviin ja web-pohjaisiin.**
- **Vertailtavat editorit:**
 - TopBraid Composer (Free ja Standard), Protégé 4.3.
 - TopBraid EVN, PoolParty, Vocbench.
- **Käyttövaatimukset: käsitehierarkian näkyvyys, monikielisyys, haku termin perusteella, vakaus.**

Suosittelvat editorit

- **Suosittelavin: TopBraid EVN. Kakkossijalla: PoolParty. Koneelle ladattavista suosittelavin: TopBraid Composer Free.**
- **EVN ja PoolParty melko kalliita. Ilmainen ja avoimeen lähdekoodiin perustuva Vocbench ei toistaiseksi vakuuttanut.**
- **TopBraid EVN:n etuina päivittyvyys usealle käyttäjälle, monipuoliset hakutoiminnot. EVN:ssä ontologiasta on olemassa työversio, jota voi kommentoida ennen hyväksymistä.**

Ontologiaeditorit

Editori	Tyyppi	Helppo- käyttöi- syys	Perus- toimin- not	Lisätoi- minnot	Luotetta- vuus	Hinta	Lähdekoodi
TopBraid Free	Paik.	4	4	3	4	ilmainen	suljettu
TopBraid Standard	Paik.	4	4	4	4	edullinen	suljettu
Protégé	Paik.	3	3	3	2	ilmainen	avoin
PoolParty	Web	4	4	3	3	kohtalainen	suljettu
TopBraid EVN	Web	4	5	4	4	korkea	suljettu*
VocBench	Web	3	2	3	2	ilmainen	toistaiseksi suljettu

Ontologioiden ylläpito

- Paikallisesti asennettava editori sisältää usein monipuolisesti toimintoja, mutta usea käyttäjä ei voi editoida samaa tiedostoa yhtä aikaa.
- Web-pohjainen editori mahdollistaa usean käyttäjän yhtäaikaisen editoinnin.
- Web-pohjaiselle editorille tarvitaan vain yksi asennus virtuaalipalvelimelle. Eri organisaatiot voivat kytkeytyä palvelimeen. Editorin käyttö tapahtuu selaimen kautta.

Editorien vertailuraportti

- Raportti löytyy ONKI-projektin wikin kautta.
<https://wiki.helsinki.fi/display/ONKI/>
- Julkaisun pysyvä osoite on:
<http://urn.fi/URN:NBN:fi-fe201306244091>

Sidosryhmien haastattelut

Tuomas Palonen

Tapaamiset

- Tähän mennessä haastatellut: Svenska YLE, YLEn Elävä arkisto, YLEn kirjastopalvelu, Brages Pressarkiv, Kirjasampo, THL, VNAS-ryhmä, SKS, FIRE-tutkijat.
- Tulevia tapaamisia: Viikin tiedekirjasto, JUPO-ontologian kehittäjät, Eero Hyvönen Aalto-yliopistosta.
- Erikoissanastojen ja erikoisontologioiden ylläpitäjät, tietopalveluasiantuntijat.
- Haastatteluista tehdään kooste, johon kirjataan sidosryhmien toiveita, ehdotuksia, mielipiteitä.

Esiinnousseita teemoja

- **Palvelun vakaus ja jatkuvuus.**
- **Päivittyvyys, käsitteiden ehdotusjärjestelmä.**
- **Ontologian mahdollisuuksia ei vielä hyödynnetä.**
- **Päätöksenteon haasteet organisaatiossa.**
 - Jo käytössä olevat esimerkkisovellukset auttavat
- **Oman sanaston ylläpidon tulevaisuus.**
- **Mietteitä:**
 - Jaettujen käsitteiden tarpeellisuus
 - Oman erikoisalan sanastoa puuttuu YSOsta
 - Jotkin hierarkiaratkaisut koetaan vieraiksi

Lisää esiinnousseita teemoja

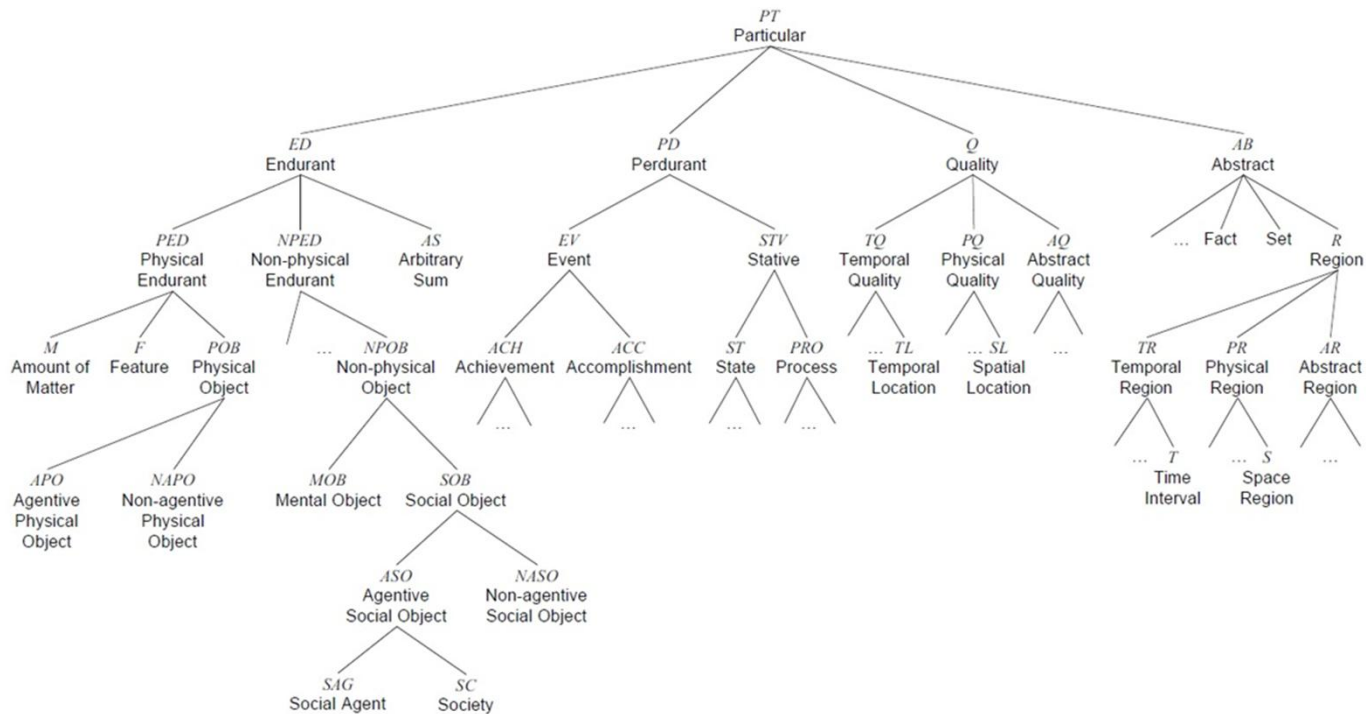
- **Assosiaatiot, horisontaalisuus**
- **Paikkaontologiat**
- **Toimijaontologiat**
- **Aineistojen linkittyvyys muualle**
- **Ontologian edut:**
 - valmis integroitu sanasto
 - autocomplete-ehdotus
 - hierarkia vähentää päällekkäistä sisällönkuvailua ja helpottaa hakua

Ehdotus Yleisen suomalaisen ontologian (YSO) uudeksi ylärakenteeksi

Mikko Lappalainen

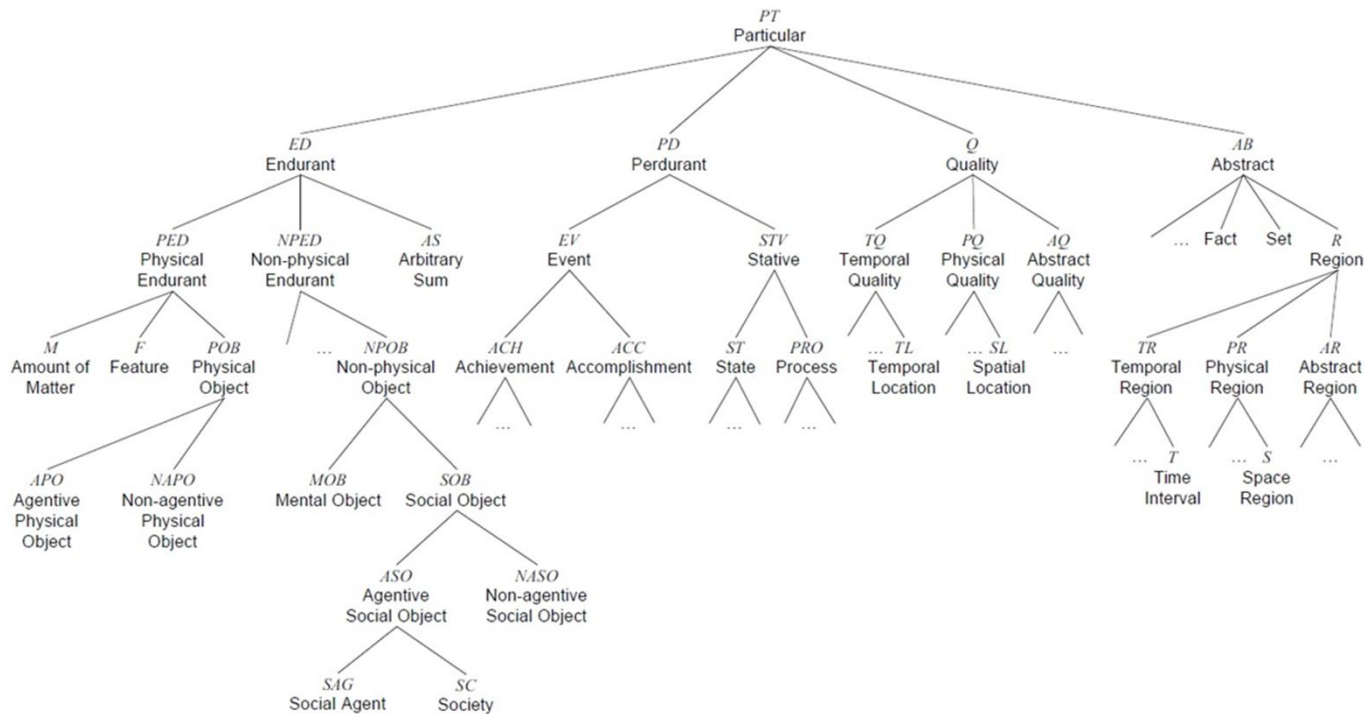
YSON ylärakenne ja FinnONTO

- YSON ylärakenne (hierarkian ylimmät tasot) on luotu FinnONTO-hankkeen alussa DOLCE-mallin pohjalta
- Ylätason jako: Endurant-Perdurant-Quality-Abstract



YSON ylärakenne ja FinnONTO

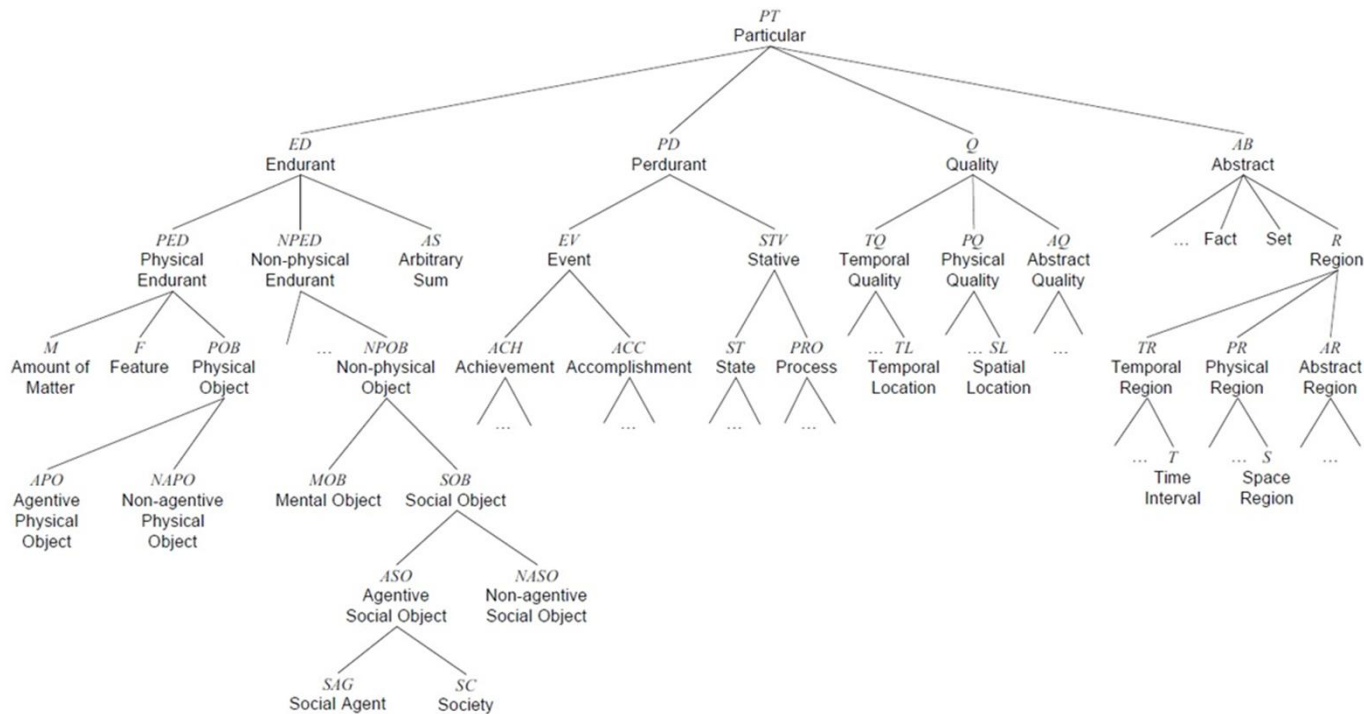
- DOLCEn Endurant-Perdurant –jako tarkoittaa käytännössä jakoa *objektien ja ilmiöiden (tai tapahtumien)* välille
- YSOssa käännetty ”Pysyväksi” (Endurant) ja ”Muuttuvaksi” (Perdurant)



YSON ylärakenne ja FinnONTO

Ongelmakohtia:

- Termit Endurant ja Perdurant filosofisesti sovittamattomia
- Käännökset Pysyvä ja Muuttuva epätarkkoja ja vaikeasti hahmotettavia
- Dolcen käytössä ei olla oltu johdonmukaisia



YSON ylärakenteeseen tarvittavat muutokset

1. ”muuttuva” ja ”pysyvä” tulee nimetä uudelleen ja niiden määritelmät tulee kirjoittaa uusiksi
2. ”ilmiöt”- sekä ”ajanjakso”-käsitteet tulee siirtää toiseen päähaaraan (DOLCEn mukaisesti)
3. ”ominaisuudet” tulee erottaa omaksi päähaarakseen (DOLCEn mukaisesti)

YSON ylärakenteeseen tarvittavat muutokset

Miksi puuttua ongelmakohtiin?

1. Lähestymme DOLCEn alkuperäistä mallia

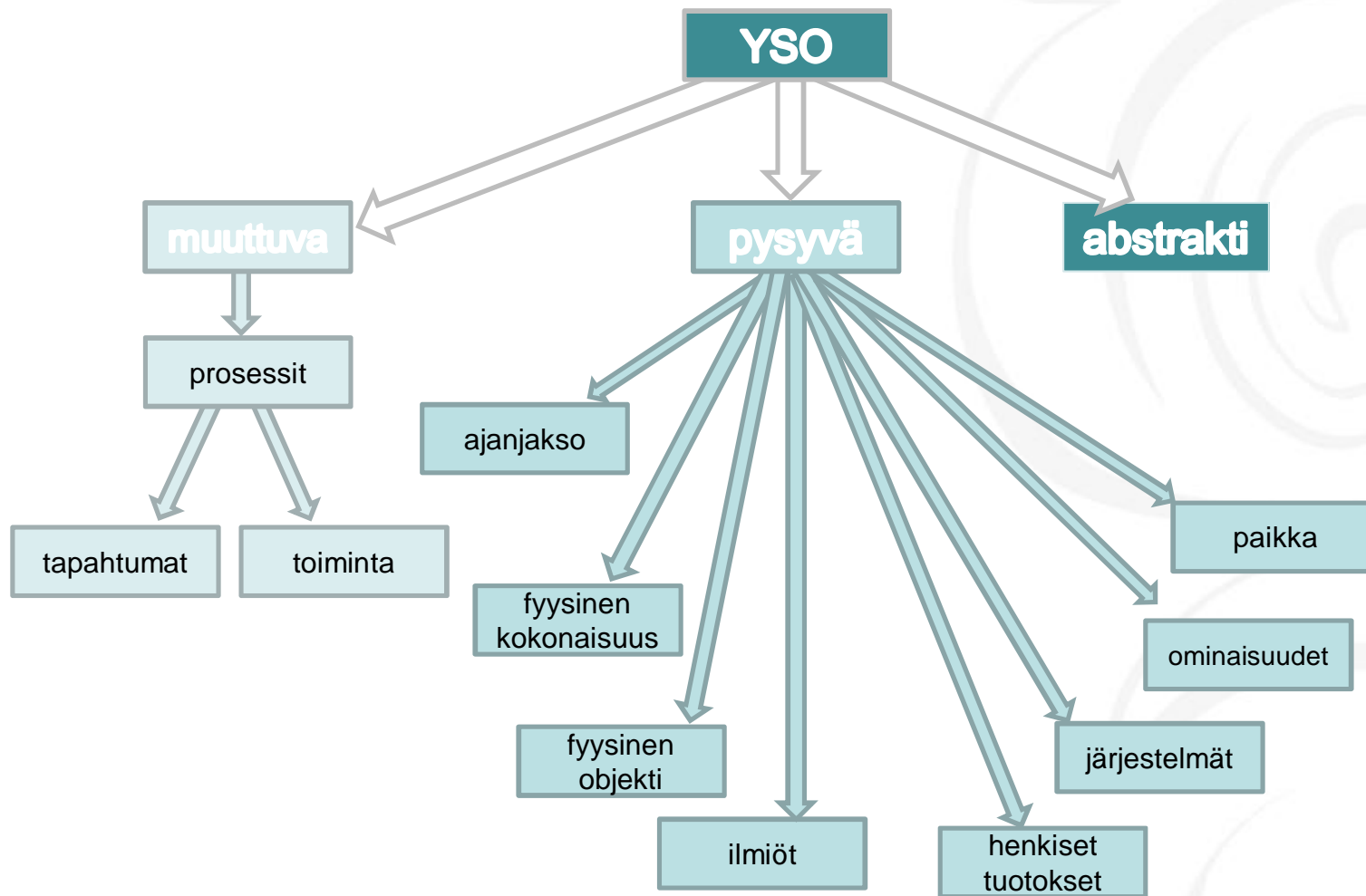
- peilaaminen kansainvälisiin ontologioihin helpompaa, DOLCE mäpätty esim. WordNetiin

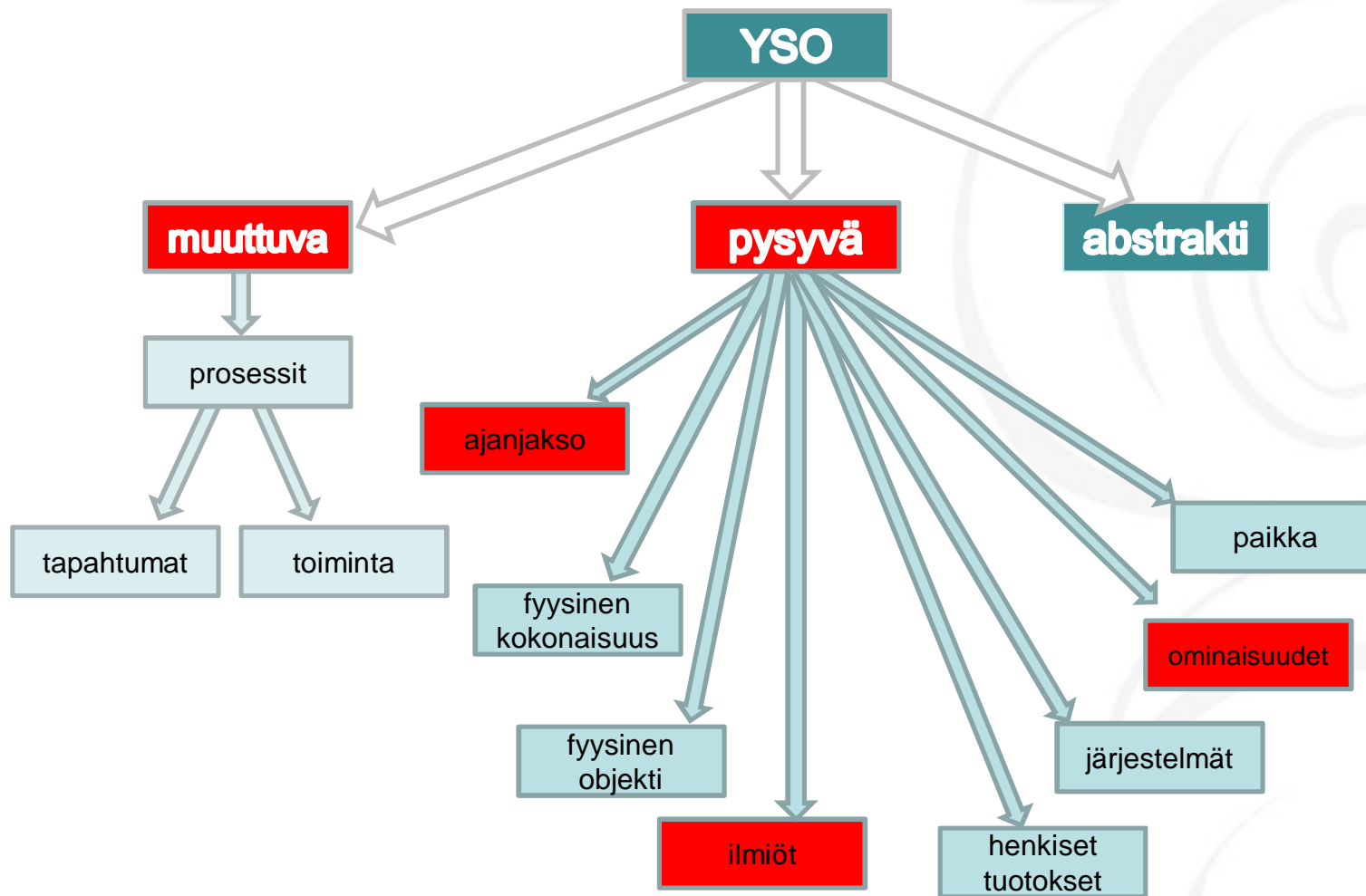
2. Otamme käyttöön kolmijaon (tapahtumat-objektit-ominaisuudet), jonka analyysillä on hyvin pitkät perinteet metafysiikan piirissä

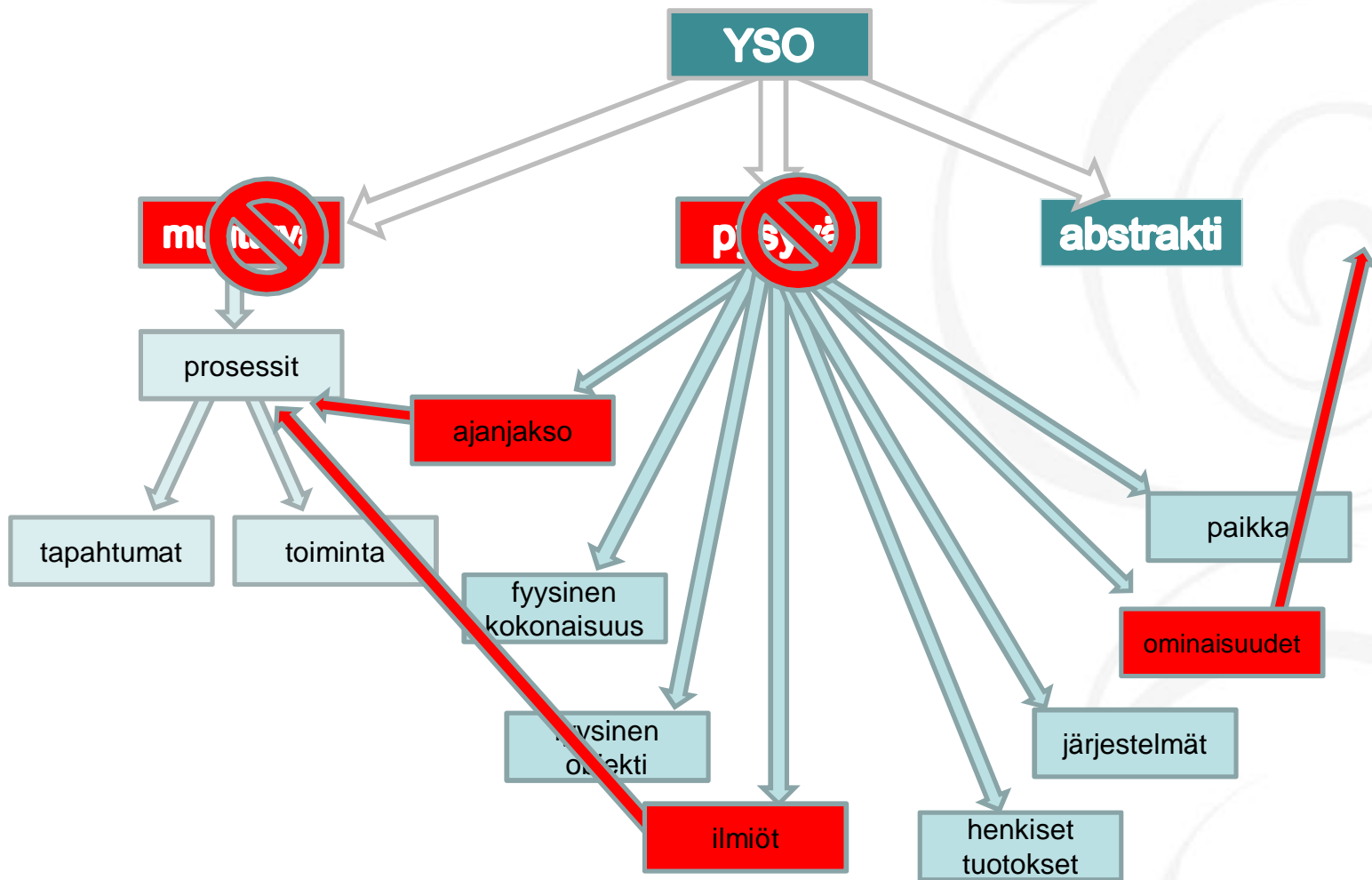
3. Kun yläkäsitteille luodaan selkeät määritelmät, jotka niiden alle sijoitettavien käsitteiden tulee täyttää, on koherentin ontologian rakentaminen huomattavasti helpompaa

- periaatteet sekä selkeä alkuajatuksen mukainen jako mahdollistavat tulevaisuudessa "loogisen päättelyn" ja ominaisuuksien periyttämisen eri päähaarojen sisällä, seikat joiden mahdollistajana ontologiat usein mainitaan

4. Hajautetun ontologiakokonaisuuden rakentaminen ja ylläpito helpompaa, kun ylärakenne on vahvalla pohjalla ja hyvin perusteltu







YSO

Tapahtumat
ja toiminta

Oliot

abstrakti

Ominaisuudet

prosessit

tapahtumat

ajanjakso

toiminta

ilmiöt

fyysinen
kokonaisuus

fyysinen
objekti

henkiset
tuotokset

paikka

järjestelmät

YSON uusi ylärakenne

- **TAPAHTUMAT JA TOIMINTA** (eng. Events):
 - 1.yhteys kausaliteettiin: jokin syy ja/tai seuraus, esim. vallankumouksella syy ja seuraus, autolla (objekti) ei niinkään
 - 2.ilmiöt "tapahtuvat" (happen, take place, occur), ovat vain osittain läsnä tietynä hetkenä
 - 3.suhteellisen selvät temporaaliset rajat, mutta epäselvät spatiaaliset rajat
 - 4.voivat olla olemassa päällekkäin samassa paikassa
- **OLIOT** (Objects, Substance):
 1. täyttävät useimmiten spatio-temporaalisen tilan
 - 2.objektit "ovat olemassa" (they exist), läsnä ajassa kokonaisuuksina
 - 3.suhteellisen selvät spatiaaliset rajat, mutta epäselvät temporaaliset rajat
 - 4.sijaitsevat paikassa itsenäisinä; eivät siedä päällekkäisyyttä
- **OMINAISUUDET** (Properties):
 - luonnehtivat (characterize) objekteja ja ilmiöitä, esim. muoto, väri, paino, pituus
 - "realisoituvat" objekteissa tai ilmiöissä
 - ominaisuuden instanssi (esim. tietty punaisuuden sävy) voi olla läsnä useassa objektissa (vrt. tietyn objektin instanssi esim. Sokrates uniikki)

YSON uusi ylärakenne

Muutosten hyötyjä:

1.Lähestymme DOLCEn alkuperäistä mallia

- peilaaminen kansainvälisiin ontologioihin helpompaa, DOLCE mäpätty esim. WordNetiin

2.Otamme käyttöön kolmijaon (tapahtumat-objektit-ominaisuudet), jonka analyysillä on hyvin pitkät perinteet metafysiikan piirissä

3.Kun yläkäsitteille luodaan selkeät määritelmät, jotka niiden alle sijoitettavien käsitteiden tulee täyttää, on koherentin ontologian rakentaminen huomattavasti helpompaa

- periaatteet sekä selkeä alkuajatuksen mukainen jako mahdollistavat tulevaisuudessa ”loogisen päättelyn” ja ominaisuuksien periyttämisen eri päähaarojen sisällä, seikat joiden mahdollistajana ontologiat usein mainitaan

4.Hajautetun ontologiakokonaisuuden rakentaminen ja ylläpito helpompaa, kun ylärakenne on vahvalla pohjalla ja hyvin perusteltu