

Kansalliskirjaston ONKI- palveluun otettavat sanastot ja ontologiat

Osma Suominen

21.10.2013

ONKI-hankkeen laajennettu projektiryhmä



Sisälllys

1. ONKI3:ssa julkaistut sanastot ja ontologiat
2. Ehdotus Kansalliskirjaston palvelun sanastoiksi ja ontologioiksi
3. Käytännöt uusille sanastoille

ONKI3:n sanastot ja ontologiat

Aiheen mukaan

- 9 Yleiskäsitteet +
- 19 Alakohtaiset +
- 5 Julkishallinto +
- 13 Kulttuuri +
- 7 Liike-elämä +
- 14 Luonto +
- 6 Paikkatieto +
- 10 Terveys +

Tyypin mukaan

- 35 Luokkaontologia +
- 18 Yksilöontologia +
- 25 Kehittynyt sanasto +
- 5 Yksinkertainen sanasto +

Julkaisutavan mukaan

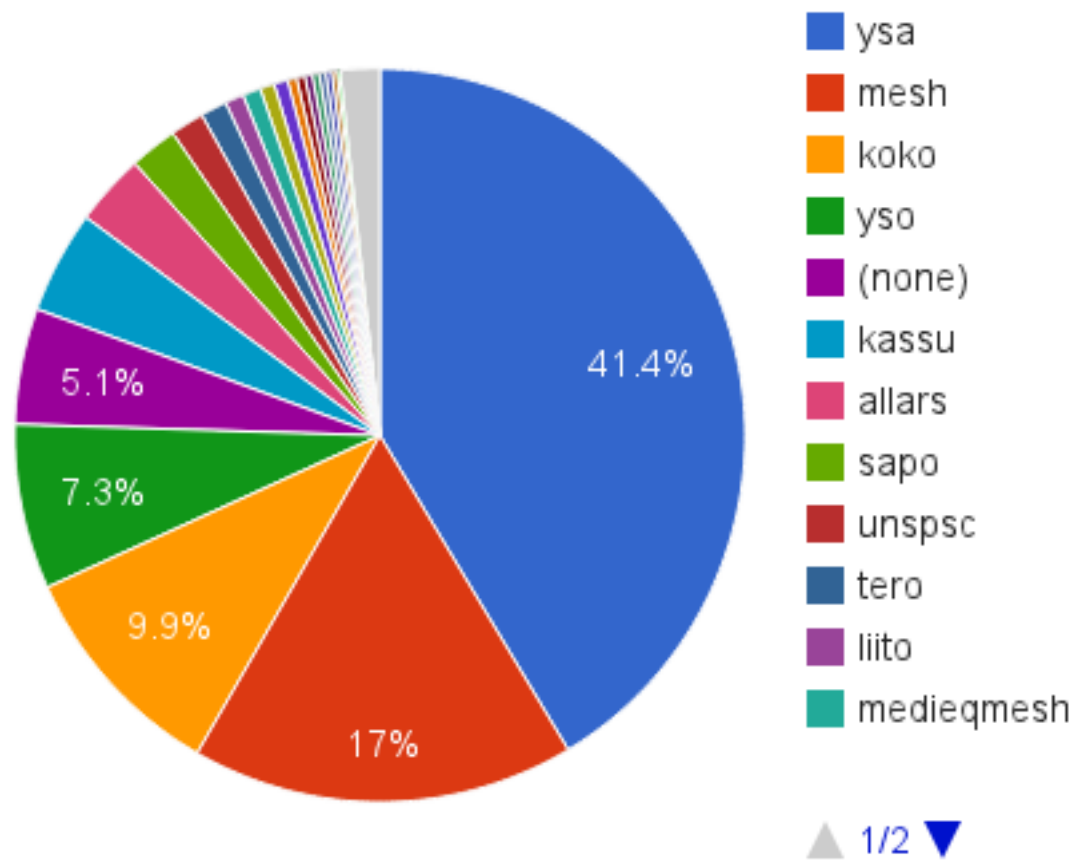
- 48 Julkiset ONKI-ontologiat +
- 18 Yhdistelmäontologiat kehittäjille +
- 5 Julkiset OMA ONKI -ontologiat +
- 12 Asiakaskohtaiset ontologiat +

Julkaisijan mukaan

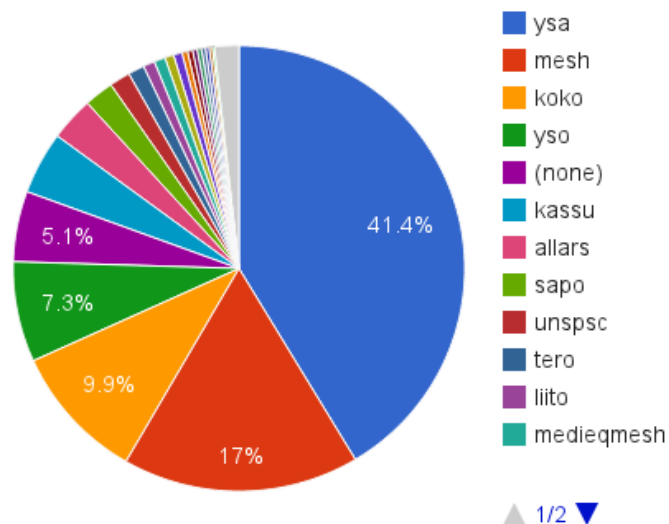
- 3 Kansalliskirjasto +
- 30 FinnONTO-konsortio +
- 1 Aalto-yliopiston taiteiden ja suunnittelun +
- 1 BirdLife +
- 1 Brages Pressarkiv +
- 1 Helsingin yliopiston kirjasto Viikin +
- 1 Kirjastot.fi +

- ▶ **Aikakauslehtisanasto** ↻
- ▶ **Allmän tesaurus på svenska - Allärs** ↻
- ▶ **Creative Commons Licenses** ↻
- ▶ **DCMI Type Vocabulary** ↻
- ▶ **Disaster ontology** ↻
- ▶ **Erytisryhmä-ontologia**
- ▶ **Eurooppalainen monikielinen terveyden edistämisen sanasto - HPMULTI** ↻
- ▶ **Iconclass**
- ▶ **Integrated Public Sector Vocabulary - IPSV** ↻
- ▶ **IPTC Genre** ↻
- ▶ **IPTC Media Topic** ↻
- ▶ **IPTC Scene** ↻
- ▶ **IPTC Subject Code** ↻
- ▶ **IPTC Subject Qualifier** ↻
- ▶ **IPTC World Region** ↻
- ▶ **Julkishallinnon ontologia - JUHO** ↻
- ▶ **Kassu - kasvien suomenkieliset nimet** ↻
- ▶ **Kaunokki-ontologia - KAUNO** ↻
- ▶ **Kielitieteen asiasanasto**
- ▶ **Kielitieteen ontologia - KTO** ↻
- ▶ **Kielten nimien ontologia** ↻

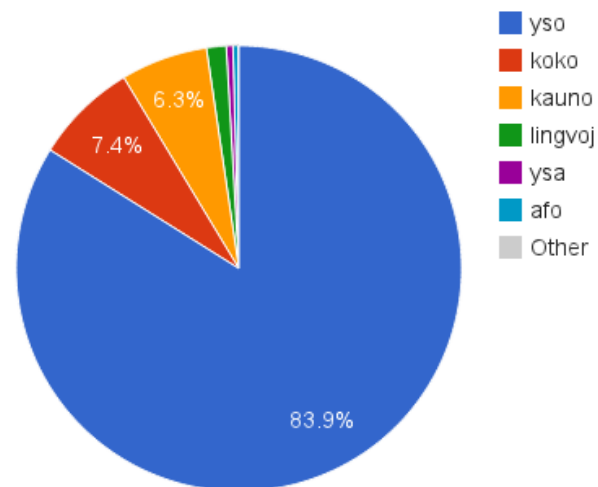
ONKI3 käyttöliittymällä käytetyt sanastot



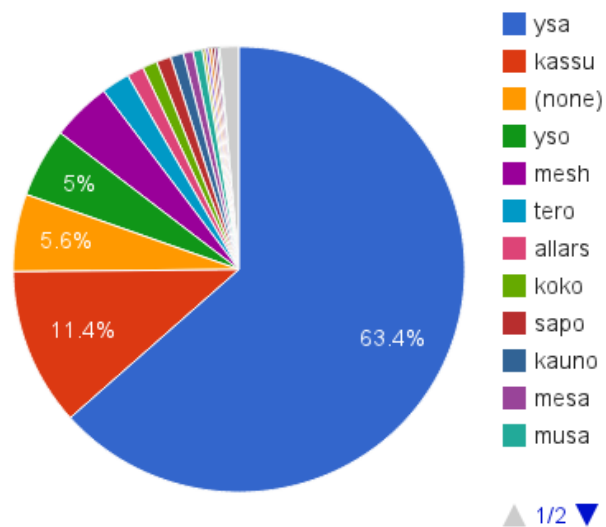
ONKI3 käyttöliittymällä käytetyt sanastot



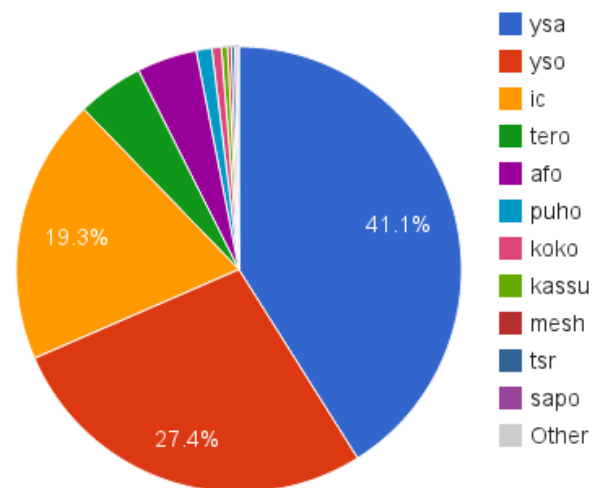
SOAP rajapintakäyttö sanastoittain



HTTP rajapinnan käyttö sanastoittain



DWR-rajapinnan (ONKI-valitsimen) käyttö per sanasto



ONKI3:n sanastot: yhteenveto

- paljon sanastoja
 - 83 näkyvillä + muutama piilossa
- käyttö keskittyy melko pieneen joukkoon
 - long tail -ilmiö
 - vaihtelee hiukan käyttöliittymän/rajapinnan mukaan

Ehdotus Kansalliskirjaston palvelun sanastoiksi ja ontologioiksi

Tavoitteet

Julkaistaan sanastot ja ontologiat, jotka

- kattavat valtaosan nykyisestä käytöstä ja/tai
- vastaavat tulevaisuuden tarpeisiin

Lisäksi KOKOn osaontologiat otetaan mukaan eheyden säilyttämiseksi, vaikka monia niistä ei juuri käytetä sellaisinaan.

Ehdotus taulukkona:

<http://tinyurl.com/kk-sanastot>

YSO, YSA, Allärs, MUSA/CILLA

- Kansalliskirjaston ylläpitämiä
- Päivitetään säännöllisesti
- Voidaan julkaista aina tuorein versio
 - esim. automaattipäivitys joka yö

KOKO + erikoisontologiat

- Jo nyt tarjolla KOKOn “betaversio”, jossa mukana uudehko YSO ja erikoisontologiat
- Suunnitelmissa kehittää KOKOa ensi vuodesta alkaen
- Mukana myös kaikki KOKOn osaontologiat
 - JUHOn ja JUPOn (tekninen) ylläpito mainittu KK:n ONKI-hankesuunnitelmassa
 - muut julkaistaan eheyden nimissä, vaikka monia niistä ei tällä hetkellä ylläpidetä

Paikkaontologiat

- vanhastaan olemassa SAPO ja SUO, joita ei ole ylläpidetty (SUO 2008, SAPO 2010)
- SAPOa kysytty paljon
 - mahdollinen jatkoylläpito SeCo-tutkimusryhmässä
- Maanmittauslaitoksen Paikannimirekisteri
 - julkaistu avoimena datana 2011
 - voidaan julkaista merkittävimmät paikat (esim. 60000 kpl) suoraviivaisena SKOS-muunnoksena
 - korvaa SUOn (suomalaisten paikkojen osalta)

Iconclass

- ollut tarjolla ONKI3:ssa jo pitkään
- julkaistu iconclass.org:issa avoimena linkitettyinä datana vuonna 2012
- voidaan julkaista uusi versio, mutta URI-tunnisteet muuttuvat
 - esim. “kuningas”:
vanha: http://seco.tkk.fi/onto/iconclass/iconclass#not_44B113
uusi: <http://iconclass.org/44B113> (loppuosa pysyy samana)

Lingvoj ja Lexvo

- ONKI3:ssa ollut tarjolla Lingvoj-sanasto joka luetteloi koko maailman kielet ja kielikoodit.
- Lingvoj suosittelee nykyään Lexvon käyttöä.
 - Lexvo on uudempi vastaava luettelo
 - Muunnos: <http://www.lingvoj.org/lingvoj-to-lexvo.rdf>

Käytännöt uusille sanastoille

Kriteerit julkaisulle (ehdotus)

1. Pyyntö tulee ONKI-konsortion jäseneltä
 - ei julkaista sanastoja jos ei ole tiedossa käyttöä
2. Ylläpito järjestetty
 - ei haluta kuolleita sanastoja (ilman hyvää syytä)
3. Ei kohtuutonta vaivaa palvelun ylläpidolle
 - optimitapauksessa valmiiksi SKOS-muodossa
 - suoraviivaiset tekniset muunnokset mahdollisia
4. Käyttöehdot sallivat avoimen julkaisun

Kiitos

tämä esitys: <http://tinyurl.com/kk-sanastot-esitys>