

YSO-Wikidata - linkitys

Joeli Takala

YSO-Wikidata - linkitys

- YSO on varsin kattava ja huolellisesti ylläpidetty ontologia, mutta Suomen ulkopuolelta katsottuna vain yksi kansallinen ratkaisu muiden joukossa
- YSO:lla kuvailtu aineisto linkittyy lähinnä YSO:lla, KOKO:lla tai jollain KOKO:on linkitetyllä erityisontologialla kuvailtuun aineistoon
 - Etsitään käsiimme jokin sellainen kansainvälisesti merkittävä tietämuskanta, johon muutkin yleisontologiat linkittyvät

YSO-Wikidata - linkitys

- Päätettiin käyttää YSO:ssa olemassa olevia LCSH-linkkejä, jota kautta etsittiin yhteys Wikidataan
- Wikidata on vapaasti käytössä oleva tietämuskanta joka on sekä ihmisten että koneiden muokattavissa ja luettavissa
- Wikidatan koko on reippaassa kasvussa: 15M käsitettä 2016, 56M käsitettä 2018
 - Google vienyt Freebase-ontologian Wikidataan
- YSO-paikoille tätä linkitystä on jo tehty!



YSO-Wikidata - linkitys

- Valittiin datasetiksi Finnassa yleisimmin käytetyt YSA-käsitteet yhdistettynä YLE:n tekemiin linkityksiin KOKO-käsitteiden ja Wikidata-käsitteiden välillä
→ saatiin 5824 YSO-käsitettä joista puolille on suoraan linkki Wikidataan
- Verrattiin YSO:sta löytyviä LCSH-linkkejä Wikidatasta löytyviin
→ saatiin Wikidata-linkit 82%:lle kysytyistä YSO-käsitteistä
- Tulosten tueksi haettiin vielä merkkijonohakuna mahdolliset linkit Wikidatan sisältöihin jokaiselle YSO-käsitteelle linkitysten tueksi käsityönä tehtävää tarkistusvaihetta varten
→ Näitä kolmea tietolähdettä ristiinvertaamalla saadaan linkit kaikille YSO-käsitteille Wikidataan

YSO

Wikidata

LCSH

YSO-Wikidata - linkitys

- Tulokset julkaistaan Wikidatassa (eikä pelkästään YSO:ssa), jolloin YSO toimii aidosti monikielisenä tietämuskantana
→ YSO:n tietorakenteisiin pääsee käsiksi helpommin, Wikidatan kautta
- Wikidatalla löytyy valmiit työkalut linkittämiseen
→ YSO-Wikidatayhteys on myös automaattisesti ylläpidettävissä
- Jos käsintarkisteluvaiheessa automaattisesti generoidut linkit osoittautuvat tarpeeksi hyväiksi, voidaan periaatteessa sama linkitys ajaa koko YSO:lle. Koesetti kattaa noin 20% YSO:sta.