

2018-08-31 YSAsta YSOon konversioryhmän kokous 4

Aika: Perjantai 31.8.2018 klo 13.00-15.30

Paikka: Kansalliskirjasto, Fabiania, kokoushuone Pohto

Etäyhteys: https://connect.funet.fi/kk_finto/

Osallistujat

Paikalla	Etänä	Poissa
- Mikko Lappalainen, KK, pj. - Tuomas Palonen, KK - Matias Frosterus, KK, siht. - Alex Kourijoki, KK - Jarmo Saarikko, KK - Mirja Anttila, KK - Ulla Ikäheimo, KK - Tarja Mäkinen, KK - Jaakko Tuohiniemi, HULib - Ville Huhtala, HULib - Maria Kovero, HULib	- Maria Forsén, ÅAB - Saijamari Pakkala, PIKI - Anne Viitanen, Vaski - Nicholas Volk, KK	Minttu Hurme, KK

Edellinen kokous: [2018-06-06 YSAsta YSOon konversioryhmän kokous - ketjut](#)

Huomioita

- Joitain (auktorisoiattomia) ketjuja vastaa eksplisiittisesti yksi YSO-käsite. Nämä ketjut pitäisi käsitellä eri tavalla kuin "tavalliset" ketjut - ne pitäisi tunnistaa ja korvata yhdellä YSO-käsitteellä
- Konversio nostaa esiin laatuongelmia, mutta näiden korjaus ei periaatteessa ole tämän projektin kontolla
- Kuka päättää siitä puretaanko ketjutus vai ei?
 - KUMEA ja Sisällönkuvailuverkosto ohjaa Melindan kuvailua - voisivatko KUMEA ja Sisku antaa suosituksen ketjujen purkamisesta?
 - Melinda on laaja yhteistyöverkosto ja tämän tyyppiset asiat eivät välttämättä ole yhden tahon päätettävissä.
 - Ketjujen luominen voidaan estää teknisesti (kohdistetusti YSO/Allfoon), jolloin uusien ketjujen syntyminen voidaan estää
 - Vai voidaan? Entä uusilta organisaatioilta tulevat ketjut? Ja kun Aleph vaihtuu, voivat estoviritykset hajota samalla
 - Entä Kaunokin ketjut ja ketjuttamismahdollisuus?
- Vain YSA-termit konvertoidaan
 - ja Allärs
 - ja MUSA/Cilla
 - YSOon tai SLM:ään kentästä riippuen
- Violassa kannattaisi tehdä MUSA->YSA/YSO-konversio ennen Violan liittämistä Melindaan
 - laajemminkin kaikki Melindaan tulevat tietokannat kannattaisi konvertoida ennen liittoa
 - MUSA-konversiota ei voi tehdä ennen kuin Finna tukee hakuja
 - Mikä on marssijärjestys? MUSAYSA ja sitten YSAYSO vai YSAYSO ja sitten MUSAYSO
 - vai tehdäänkö ensin kirja-aineistot MUSAYSAYSO ja sitten vasta myöhemmin nuotti- ja av-aineistot (joiden MUSA-termit vaativat kenttäsiirtoja ja muuta erikoiskohtelua)
- Tästä kokouksesta pitäisi koostaa lista asioista, jotka koskevat Violaa, ja sitten tämä lista toimitetaan Katrille
 - otettava huomioon Viola-Melinda-yhteydessä
- KK tekee myös paikalliskantojen konversiot
 - KK pääsee Voyager-paikalliskantojen palvelimille, joissa konversion voi tehdä
 - Muut järjestelmätoimittajat ajavat konversio-ohjelman itsenäisesti
- Demo Melinda-testikannasta Aleph-luetteloinnista
 - Huomioita:
 - Sanastovalintaikkuna hidastaa liikaa
 - pitäisikö englanninkielisten termien poimiminen estää?
 - vai pitäisikö enPrefLabelit muuttaa enAltLabeliksi MARC-versioon
 - kopioidaan kaikki suomenkieliset ruotsinkielisiksi ja toisinpäin
 - pitäisikö konversio-ohjelman tuottaa kolme vaihtoehtoa (ruotsi, suomi ja ruotsiPlusSuomi)
 - ohjaustermit, jotka ohjaavat moneen eri termiin (eli että kahdella käsitteellä on sama altLabel) on haastava ohjelman kannalta
 - kenttien järjestys tsekattava: joka paikassa missä on i pitää olla i ensin paitsi jos on w**
 - järjestys sellainen, että laajemmat, suppeammat ja sitten rinnakkaiset ja aakkostettu jokaisessa**
 - Kuinka paljon yksiköt ja monikot menevät päällekkäin muiden auktorisoitujen juttujen kanssa?
 - PÄÄTÖS:
 - ei mahdollisteta englanninkielistä kuvailua
 - toimitetaan MARC-tietueet kaikille
 - otetaan linkit mukaan tietueisiin
 - samaten koordinaatit, jähka varmistetaan kenttä
- Seuraava kokous 2.10. klo 9.30-12
- Kaikki päätökset voisi laittaa aswiin
- Lisää rummutusta siitä, että konversio on tosiaan tulossa

Asialista

- Kerrataan edellisen kokouksen päätökset (Jarmo)
- Konversiosuunnittelun eteneminen ja syksyn työn aikataulutus (Mikko)
- Violan ja Melindan **musa/cilla** -termien konversiot (Mikko)
- Ajankohtaiskatsaus: mitä tehty kesän aikana mitä tehdään jatkossa (Alex)
- Ketjutus part 2 (Mikko)
- Wikidata-linkit (Mikko)
- Konversiosääntöjen valmistelu (Mikko)
- YSO:n käyttö kuvailussa : demo Melindan testipuolella (Jarmo)

Suunnittelun aikataulu

Konversior ryhmän kokoukset syksyn aikana

- Montako, aiheet, milloin?
- Missä järjestyksessä konversiosääntöjä ryhdytään käsittelemään, jotta saadaan valmista ennen vuoden loppua?

PÄÄTÖKSET:

Musa+Cilla2yso konversio

Kuka hoitaa? Tuleeko tehdä ennen ysa-yso -konversiota?

Tapanilta ja Jaakolta saatu vastauksia

- Muusan vastuulla on sisällönkuvailuun liittyvän ohjeistuksen laatiminen musiikin kuvailua varten, mutta varsinainen konversio ei sille kuulu.
- Melindan osalta Ulla Ikäheimo osanee vastata kysymykseen aikataulutuksesta ym.
- Violan puolesta voidaan todeta, että Musa-termejä on käytetty hyvin niukasti aihe termeinä musiikkiaineiston kuvailussa.
- Musa/Cilla -termit 650 kentässä
 - Musa–Ysa-konversio lieene tässä kokonaisuudessa se helpoin osuus.
 - Luomisaika, luomispaikka, esityskokoonpano jne. siirretään ensin asianmukaisiin kenttiin, jotta jäljelle jäävät 650-kenttien MUSA-termit voitaisiin sitten sitten heittää helpommin YSA- tai YSO-termeiksi
- Musa/cilla-termit 655 kentässä
 - Suurin osa Musa-termeistä on muoto- ja lajityyppitermejä, joiden muuntaminen SLM-termeiksi, uusien kenttien käyttöönotto jne. vaatiikin sitten perusteellisempaa suunnittelua. Mutta

PÄÄTÖKSET:

#ysall2yso tilannekatsaus

Elokuu 2018

- tuotettu uusi versio 0.5 konversio-ohjelmasta
 - korjattiin konversio-ohjelman edellisen ohjelman tuotoksissa havaitut virheet
 - tuotettu LSH linkit MARC-tietueisiin
 - toimitettu uudet MARCXML-tiedostot Melindan testipalvelimelle
- Auktoriteettitiedosten testaus alkanut tällä viikolla (Jarmo, Ilona, Satu)
- YSO voitaneen julkaista Asterissa mikäli testissä ei havaita ongelmia ja konversior ryhmä näyttää vihreää valoa
- Konversion ennakkotyöt: jaettujen YSA-käsitteiden termien korjauksia tehty Melindassa

Käynnissä

- yso2marc python-koodin refaktorointi (siivous ja irrottaminen Jupyter-notebookista itsenäiseksi ohjelmaksi)
- täydennetään konversio-ohjelma käsittelemään SLM sanaston käsitteet ja niiden ominaisuudet niitä vastaaviin MARC -kenttiin
- lisätään wikidata-linkit sekä **yso** että **yso-paikat** MARC tietueisiin (750/751 kenttään)
- lisätään PNR linkit **yso-paikat** MARC-tietueisiin (751 kenttään)

Jatkotoimenpiteet

- Kun **yso-paikat** -käsitteille on lisätty koordinaatit (vain pistekoordinaatit, PNR/Wikidata), ne voidaan viedään MARC-tietueisiin (034 kenttään)
- Laji- ja muototyyppisanasto (**slm**) voidaan tuottaa MARCXML-muotoon ja viedään Asteriin samanlaisella prosessilla kuin **yso** (155/455/755 kentät)
- Tavoitteena tuottaa cronista päivittäin ajettava python-skripti, joka tuottaa Finton yso:sta ja slm:stä MARCXML-tiedostot
- Tuotetut XML-tiedostot siirretään Melinda-palvelimelle ja ajetaan Asteriin (Melinda-palvelu vastansee tästä puolesta)
- XML-tiedostot asetetaan Fintoon saataville ja viedään versionhallintaan (github)
- Selvitettävänä: käsitteen määritelmä-kentän Lähde: merkinnän koodaus yso:ssa ja sijoitus MARC 670 -kenttään (tuotetaan yso:n skos:note -kenttään Voyagerin 670-kentästä)

- Konversio-ohjelman suunnittelu käynnistyy syyskuussa

PÄÄTÖKSET: Laitetaan tarjolle jonnekin, josta kuka vain saa sen. Lisäksi mietitään saisiko rajapinnan, jonka avulla saisi vain uudet ja muuttuneet.

MUSA-auktoiteettien osalta tarvitaan päätös, että Melindassa ei voi valida MUSA-termejä (poistuvat Asterista)

Tavoite: MUSA-poistetaan Asterista 1.10. mennessä

Ketjutus part 2

Miten päätös ketjujen purkamisesta jalkautetaan kentälle päättävälle tahoille?

Asiantuntijaryhmien kautta?

PÄÄTÖKSET:

Wikidata-linkit

Viedäänkö auktoiteettitietueisiin? yso: vai yso-paikat: vai molempien tietueisiin?

PÄÄTÖKSET:

Konversiosäännöt

Tekstuaalinen dokumentointi: Master dokumentin sisältö ja rakenne

- esimerkiksi aikaisempi Melindan MARC-konversion massiiviset csv-taulukot
- voisi olla google sheet, jota ryhmän jäsenet voivat seurata

Konversiosääntöjen käsittelyä varten tarvitaan esimerkkejä. Koottu tänne:

- haastavaksi tunnettuja termitapauksia KK:sta, **Jaakolta** ja **Villeltä**
- Ongelmatapauksia erilaisia ongelmatyyppejä, joita joudutaan ratkaisemaan:
 - korjaamattomat Alllärs-termit
 - jaetut käsitteet
 - sulkutarkenteet
 - ketjun eri osat
 - eri aineistotyytit
 - esim. helkassa ketjujen purussa usein havaittuja ongelmia (käsin puretut jne.)

PÄÄTÖKSET: Google-docs vaikeista tapauksista, joita kerätään seuraavaan kokoukseen

konversiosäännöt taulukoksi ("tämä kenttä siirtyy tähän kun täältä löytyy tämä")