



Melinda

Melindan massakorjauksista

Nicholas Volk

Kansalliskirjasto, Melinda-palvelut



Melindan massakorjauksista

- Melindassa on paljon hyvää metadataa, mutta seassa myös vähemmän hyvää
- Osa korjauksista on aitoja virheitä, osa parannuksia tai täydennyksiä, osa juontaa alkuperänsä muuttuneisiin luettelointikäytäntöihin yms.
- Tunnetut virheet ja muutokset on parempi ja helpompi korjata ennen linkitettyyn dataan siirtymistä
- Osa massakorjaustarpeista tunnistetaan itse, osa tulee esim. Kumealta ja osa yksittäisten kirjastojen edustajilta suoraan
- Korjauksia voi ehdottaa osoitteeseen melinda-posti@helsinki.fi

Replikoinnista

- Massakorjaukset eivät tyypillisesti siirry paikalliskantoihin
- Taustalla kirjastojen omat replikointiasetukset, ja kaikilla ei replikointia käytössä
- Muutosajot kuormittaisivat paikalliskantaja ja saattaisivat häiritä luotteloitsijoiden työtä

Melindan massakorjaustavoista

- Suurin osa korjauksesta ajetaan manuaalisesti
 - poimitaan relevanttien id:iden joukko Melindasta
 - ohjelmoidaan korjaus ja testataan se
 - ajetaan varsinainen korjausajo
- Osa korjauksista ajetaan automaattisesti iltaisin
 - yöllä otetusta tietuedumpista poimitaan ns. virhesnifferillä mielenkiintoiset tapaukset ja korjataan ne
 - esimerkiksi asiasanojen käännös tehdään näin

Läpinäkyvä joukkoistusyritys ennen korjausesimerkkejä

- Virhesnifferi löytää paljon enemmän virheitä/ongelmia kuin osataan korjata automaattisesti
- Jos joku haluaa dumpin ohjelman löytämistä oman kirjaston tietueissa olevista ongelmista, niin voin pyynnöstä poimia

020\$q: sidotun ja nidotun konvertointi kovakantiseksi ja pehmeäkantiseksi

- Kumean alkuperäinen pyyntö koski kenttää 020. Tehtiin myös kentille 015, 024 ja 028.
- Sidosasutermi konversio kansityyppitermeiksi tehtiin oheisen taulukon mukaisesti. Aiemmin on jo konvertoitu rengaskirjat kierreseliksi ja kaksiosaiset termit, kuten nidottu, kovakantinen kovakantisiksi.
- Konversiota saatu Melindassa valmiiksi marraskuussa (pl. pienet loppusiivouset). Paikalliskantoja kannustetaan konversion tekemisen.

Nykyinen 020\$q	konvertoitu 020\$q
sidottu	kovakantinen
nidottu	pehmeäkantinen
inbunden	hårda pärmar
häftad	mjuka pärmar

Tulossa pian: Fennican lisäpainostietojen siirto

- Vrt. Kumean uutiskirje tänään
- Alkutilanne:
 - 500 ## \$a Lisäpainokset: 3. painos 2017 - 6. painos 2020.
 - 500 ## \$a Lisäpainokset: 6. painos 2020. ‡9 FENNI<KEEP>
- Välissä:
 - 500 ## \$a Lisäpainokset: 3. painos 2017 - 6. painos 2020.
 - 500 ## \$a Lisäpainokset: 6. painos 2020.
 - 594 ## ‡a Lisäpainokset kansalliskokoelmassa: 6. painos 2020 ‡5 FENNI
- Loppusiivon jälkeen
 - 500 ## \$a Lisäpainokset: 3. painos 2017. - 6. painos 2020.
 - 594 ## ‡a Lisäpainokset kansalliskokoelmassa: 6. painos 2020 ‡5 FENNI

518: tapahtuman aika ja paikka -huomautus

- Kentässä 518 oli usein kaikki tiedot laitettu osakenttään \$a:

518 \$a Äänitys: 1970, Jyväskylä, Taulumäen kirkko.

- Näiden rakennetta on yritetty tunnistaa ja laittaa osat informatiivisempiin osakenttiin osakenttiin:

518 \$o Äänitys: \$d 1970, \$p Jyväskylä, Taulumäen kirkko.

- Koska \$a:n arvoissa oli paljon variaatiota, niin kaikkea ei osattu säännöillä korjata järkevällä työmäärällä

Disney, Walt kirjaukset 100- ja 700-kentissä

- Toimija, jolla on näihin aineistoihin jokin suhde, on Walt Disney Company, ei henkilö nimeltä Walt Disney.
- Kansallisbibliografiassa Disney, Walt oli tekijänä n. 2 800 suomennetussa kirjassa. Melindassa tekijähaku Disney, Walt n. 7 000 osumaa.
- Korjattu:
 - 100- ja 700-kenttien Disney, Walt -kirjaukset pois
 - 245-kentissä muutettiin ensimmäinen indikaattori nolllaksi, mikäli 100-kenttä pois
 - 500-kenttään lisättiin *Walt Disney -tuotantoa*
- 1930-luvulla ja aikaisemmin julkaistut julkaisut jätettiin korjausajon ulkopuolelle

Pienempiä esimerkkejä

- 505-kentän erottimet normalisoitu " -- " vs " ; " (preferoitiin " -- "-muotoa)
- Menossa: 337 \$a oförmedlad \$b n \$2 rdamedia => \$a omedierad ...
- Jos kentässä on osakenttä \$5, niin \$9 FOO<DROP> -muotoinen osakenttä pois
- poistettu tai korjattu 084-kentästä poistettu laittomia Artosta tulleita arvoja.
- Lisäksi tulevaisuudessa 084-kentästä siivotaan muita laittomia arvoja, lähinnä tapauksia, joissa ykl-koodin perässä on ilmeisesti hyllytykseen liittyviä jonkun kirjaston tietoja:

084 L \$\$a78.66cd\$\$2ykl

084 L \$\$a78.513cd\$\$2ykl

Ykl:n mukaisista arvoista vain kirjaimet pois. Loput varmaan kokonaan pois. Ja sitten näemmä on 63000 tapausta, joista puuttu \$2...

Tulossa: Tekijäkirjausten yhdistämistä

- Taustalla Kumean toive: ”Tekijänimekekirjauksista toimijan suhderooliosakenttä \$e pois”. Koskee 600- ja 700-kenttiä
- Poistettu kutakuinkin muualta paitsi musiikista
- Musiikin puolella halutaan jakaa kenttä kahtia: normaaliksi tekijäkirjaukseksi, jossa \$e ja \$e:ttömäksi tekijänimekekirjaukseksi (yksinkertaistettu ehtoja), ks. <https://www.kiwi.fi/pages/viewpage.action?pagelId=367460674>
- Odottaa X00-kenttien yhdistimen hiomista, ja X00-kentän \$e-osakenttien sorttaamisen hiomista.

700 1# Ꞥa Sukunimi, Esa, Ꞥe sanoittaja.

700 1# Ꞥa Sukunimi, Esa, Ꞥe säveltäjä.

700 1# Ꞥa Sukunimi, Esa.

700 1# Ꞥa Sukunimi, Esa, Ꞥe säveltäjä, Ꞥe sanoittaja.



Melinda

Kiitos!

melinda-posti@helsinki.fi

Melinda-wiki: <https://www.kivi.fi/display/melinda>

<https://twitter.com/Melindapalvelu>