



Melinda

**Massakorjaukset ja tietokantahuolto
Talonmiehen tuokio 3.4.2019
Nicholas Volk**

Massamuutokset (yleistä)

- Tavoitteena minimoida käsityön määrää
- Sattuneesta syystä nyt käsitellään lähinnä asiasanojen siivousta ennen ysasta ysoon -konversiota
 - Onhan noita muitakin, esim. X00-kentissä kokoomateoksen toimittaja -> toimittaja
- Tähän hätään korjaukset on kohdistettu vain etukäteen poimittuun osajoukkoon tietueista, ei kaikkiin
 - Eli vaikka olisi kirjoitettu toimiva sääntö, niin sitä ei olla vielä välttämättä käytetty kaikkiin tapauksiin
 - Melindaan tulee koko ajan uusia tietueita ja sitä kautta uusia virheitä

Virheellisten ysa-asiasanojen käsittelyä

- Joka yö ohjelma listaa Melindasta valikoituja virheitä
 - Poimitaan virheellinen kenttä ja tietueen id
 - Mahdollinen korjausajo ajetaan vain tehokkuussyistä vain etukäteen valikoituille tietueosajoukolla
- Virheelliset (ml. vanhentuneet, ei enää sanastossa olevat) ysa-, musa- allärs- ja cilla-asiasanat ovat yksi poimittava ryhmä
- Osalle yleisimmistä virheistä voi kirjoittaa yksinkertaisia korjaussääntöjä
 - Kolme samaa kirjainta -> kaksi samaa kirjainta, jos sana löytyy asiasanastosta
 - Systemaattiset työt: "keski-aika" => "keskiaika"



Virheellinen ‡2 (lähinnä Ysa ja Allärs)

- Kannassa oli paljon sanoja, joissa on ‡2 ”valehtelee”
 - Jos sisältö ei ole ysa ja indikaattori 2:n arvo oli '0' ja asiasana näyttää englanninkieleltä, oletetaan kongressin kirjaston asiasanaksi ja poistetaan ‡2
 - Näistä muista oli paljon myös arvoa ”ysa.”
 - Yleisimpiä sanaston kirjoitusvirheitä korjattu sanastolle kuin sanastolle
 - Jos sisältö ei ole Ysan mukainen, mutta on Allärsin mukainen, vaihdetaan ‡2 allars –muotoon
 - Ja vastaavasti allars muuttuu ysaksi



Sekalaisia esimerkkejä

- 650-kentässä on valitettavasti paljon ihmisten nimiä
 - Osa poistettu, jos sama nimi löytyy 600-kentästä
 - 610 ja Asterin yhteisönimet työn alla
 - Suurinta osaa näistä ei nykysäännöillä kuitenkaan saa siivottua
- 650 ind2=7, mutta \ddagger 2 puuttuu
 - Jos Ysaa tai Allärsia, niin lisätään haluttu \ddagger 2
 - Jos ysa/allars jo käytössä jossain tietueen 650 \ddagger 2 -kentässä, niin 653:n asiasanakin voidaan muuttaa ja 650:een
 - Säännöissä piisaa reunaehtoja, joita ei kaikkia mainita tässä



Case: kirkot

- Melindassa oli tuhansia kirkot-asiasanoja, oikea muoto voi olla:
 - kirkkokunnat
 - kirkkorakennukset
 - Jotain muuta kuten seurakunta
- Listataan tulkintaa X tukevia asiasanoja ja etsitään niitä tietueesta:
 - Kirkkokunnat: dogmit, ekumenia, oppiriidat, sekularismi...
 - Kirkkorakennukset: alttarit, arkkitehtuuri, kellotapulit, tuomiokirkot...
 - Jos löytyy vain toisen tyyppisiä sanoja, niin uskalletaan muuttaa
 - Muuten käsitöitä



Tuplien ja osatuplien poistoa

- Sama ysan tai allärsin paikka (koko ketju) saattaa olla 651:ssä ja 650:ssä -> 650 pois
- Korjauksista saattaa syntyä tuplia esim. 650-kenttään -> pois
- Jos ketju eroaa vain ± 9 XXXXX<KEEP>:n osalta niin keepitön versio pois
- Jos 653:n versio on muuten sama (pl. Indikaattorit) kuin 650:n versio (ind2!=7) tai ind2=7 ja ± 2 löytyy, niin 653 pois





Melinda

Kiitos!

melinda-posti(at)helsinki.fi

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