

KITT-tilannekatsaus

Asiantuntijaseminaari 29.4.2014

Markku Laitinen, Kansalliskirjasto

markku.laitinen@helsinki.fi

<http://www.helsinki.fi/~malaitin/>

KITT2 – kadonneen tiedon metsästy

- Kirjastoilta alkoi tulla palautetta tietojen häviämisestä vanhasta järjestelmästä KITT2:een siirtymisen yhteydessä.
- Paljastui, että vanha tietokanta oli tallentanut lukuja kahteen paikkaan. "Paikassa A" oli pääosa tiedosta, yli 2 milj. tietuetta ja "paikassa B" vajaa puoli miljoonaa.
- Siirryttäessä uuteen järjestelmään uusimman dumpin perusteella on oltu siinä käsityksessä, että se sisälsi kaiken tarvittavan tiedon. Näin siirto on tehty uusimmasta dumpista suhteessa 1:1. Epäilykset dumppien erilaisuudesta heräsivät siis tarkemmassa selvittelyssä.
- Kun nuo puuttuvat luvut sijoitetaan KITT2:n nykyiseen aikasarjaan, näyttää silmämääräisesti tarkastellen siltä, että aikasarjat ovat loogiset, joten sen perusteella harkitaa "paikan B" lukujen lisäämistä KITT2:een.

KITT2 – kadonneen tiedon metsästys

Tältä se näytti...

KITT1
≈2.000.000
tietuetta

Siirto

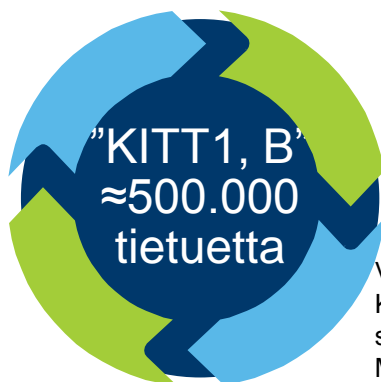
KITT2

Siirryttäessä uuteen järjestelmään uusimman KITT1:stä otetun ns. dumpin perusteella oltiin siinä käsityksessä, että se sisälsi kaiken tarvittavan tiedon, joten siirto tehtiin uusimmasta dumpista suhteessa 1:1 – kaikki näytti olevan kunnossa valtakunnan ja sektoritasolla...

Mutta...

KITT2 – kadonneen tiedon metsästyks

Tilanne olikin tämä



- Vika oli vanhassa järjestelmässä
- KITT2:ssa tätä "ominaisuutta" ei ole
- KITT2:ssa on kuitenkin vielä tehtävä laskennan tarkistuksia

Vanhemmasta dumpista löytyi "Paikka B" jossa oli käsin syötettyjä kirjastojen yläyksikköjen lukuja. KITT1:ssä, kuten KITT2:ssakin, yläyksikön lukujen puuttuessa sen luvut lasketaan alayksikköjen summana. Paikat "A" ja "B" eroavat tapauksissa, joissa yläyksikön lukuja on muokattu jälkikäteen. Muokkaus on vaikuttanut vain "Paikkaan A" ja "Paikkaan B" sisältö on pysynyt entisellään. "Paikasta B" löydettiin tietoja, joiden kohdalla "Paikassa A" oli nolla tai tähti.

Tarvitaan korjausajo – reunaehdot on...

- Ei ole mahdollista saada nyt käytössä olevia dumppeja vanhempia tietoja, niitä ei järjestelmissä enää ole.
- Siksi myös korjausajon jälkeen mahdollista, että joistakin yksiköistä puuttuu tietoja, joita on KITT1:ssä ollut.
- Suuri osa vanhalta alustalle uudelle siirtymisen yhteydessä kadonneista tiedoista voidaan kuitenkin näin palauttaa.
- **Koska vanhassa tietokannassa (KITT1) oli tietoja useammassa paikassa, on mahdollista, että jotkut tiedot ovat hävinneet jo vanhan tietokannan aikana, mutta sitä ei pysty vuosien kuluttua enää jäljittämään.**

Keskustelu etenemisjärjestyksestä

- 1 30.4.2014 klo 14: Yhteistilaston vuoden 2013 tietojen tallennus ja tarkistus päättyy ja tietokanta suljetaan syötöltä ja korjauksilta = Vuoden 2013 tilasto on valmis.
- 2 2.5.2014
 - 1 KITT2:n tuotantotietokannan kaikki tiedot siirretään testitietokantaan, johon tulee näin tuotantokannan identtinen kopio vuoteen 2013 asti.
 - 2 "paikan B" tiedot ajetaan testitietokantaan ja tiedotamme kun ajo on tehty.
- 3 3.5.2014 – **X.X.2014**: Mahdollisuus tarkistaa oman kirjastonne tiedot testitietokannassa.
- 4 **Sitä mukaa kun saatte oman kirjastonne tiedot arvioitua, lähettäkää meille palautetta tietojen oikeellisuudesta osoitteeseen kk-kitt@helsinki.fi**
- 5 **X.X.2014? Vuoden 2013 tilaston julkistaminen – MIHIN ASTI JULKISTAMISTA VOIDAAN SIIRTÄÄ, MILLOIN KIRJASTOT ALKAVAT TARVITA TIETOJA?**
- 6 Saadun palautteen perusteella arvioimme Kansalliskirjastossa hyväksytäänkö vai hylätäänkö "paikasta B" testitietokantaan tehty korjausajo.
- 7 Jos saadun palautteen perusteella korjausajon tulos on hyväksyttävä, siirrämme ajon tuloksen tuotantotietokantaan.
- 8 Vuoden 2013 tilasto julkaistaan **X.X.2014** klo 16.

Kiitos mielenkiinnostanne

Nyt on kysymysten ja keskustelun aika

Markku Laitinen
markku.laitinen@helsinki.fi
<http://www.helsinki.fi/~malaitin/>

