

Laatuongelmat ovat usein piileviä

Yhä useampi yritys tai julkishallinnon yksikkö on nykyisin "tietotalo". Talletettu tietovaranto muodostaa tärkeän resurs- sin. Tästä huolimatta tiedonhallinnassa on usein toivomisen varaa: tietoarkkitehtuuria ei ole – kenelläkään ei ole hyvää kokonais- käsitystä siitä mitä tietoja missäkin on. Tie- toja ei myöskään ole kuvattu hyvin. Kolmas ongelma on artikkelimme aiheena, nimittäin tietojen huono laatu.

Tietojen laatuongelmat ovat usein piileviä, niistä ei olla täysin tietoisia. Tilanne paljastuu yleensä tietovarastohankkeiden yhteydessä. Huomataan, että suurella vaivalla lopulta raporttiin saadut tiedot eivät olekaan julkaisukelpoisia. Piti tehdä henkilöiden sukupuolijakaumat, mutta luokassa "ei sukupuolta" onkin yli puolet henkilöistä. Kun tieto ei ollut pakollinen, ei sitä tietenkään ole syö- tetty. Henkilö, joka jaksaa näpytellä sisään kaikki tiedot ei saa palkkiota hyvästä työstä.

Toinen tilanne, missä huono laatu usein selviää, on uuden järjestelmän käyttöönottoon liittyvä konversio. Myös Master Data ja CRM –hankkeet saattavat kompuroida laatupulmien kanssa.

Laatukulttuuria tarvitaan

Osa tiedoista on nyt jo hyvälaatuista. Esimer- kiksi kirjanpitolietojen laatukulttuuri on hyvin kehittynyt. On selvää, että näiden tietojen on oltava oikein. Tilintarkastajat ovat tarkkana. Muiden tietojen osalta laatukulttuuri on lapsen- kengissä. Missä ovat asiakas-, tuote- tai henkilö- tietojen tarkastajat?

Taannoittaisessa The Data Warehouse Institut- en kyselytutkimuksessa yli 50% vastanneista oli kär- sinyt tietojen huonosta laadusta ja n. 37 % ei ollut selvittänyt asiaa. Laatua parantaneet yritykset olivat puolestaan kokeneet parempaa luottamusta analyttisiin järjestelmiin (76%), vähemmän aikahukkaa tarkistuksiin (70%) sekä päässeet paremmin "yhden totuuden" tilanteeseen (69%). Käyttäjien tyytyväisyys oli myös huomattavasti parantunut ja kuluja oli säästetty.

Vaikutuskeinoja laadun parantamiseksi

Ensinnäkin tulisi laatia yrityksen tai julkisyh- teisön laatustrategia, jossa määritellään tietojen laadulle asetettavat tavoitteet. Tavoite ei usein ole sataprosenttisesti oikeat tiedot, on määriteltävä erityyppisille tiedoille omat laatuavoitteensa. Tavoitteiden lisäksi mittarit ovat tärkeitä, niiden avulla voidaan seurata, onko tavoitteisiin päästy.

Seuraavana vaikutuskeinona ovat menetelmät, kuten tietojen profilointi, yhdistäminen ja moni- torointi.

```
Profilointi tarkoittaa erilaisia kyselyjä tietoihin, esimerkiksi:
SELECT COUNT(htun), MIN(htun), MAX(htun), AVG(htun)
FROM henkilo;
SELECT sp_koodi, COUNT(*) AS lkm
FROM henkilo, GROUP BY sp_koodi;
```

Ylempi kysely näyttää sarakkeen henkilötunnus pienimpiä ja suurimpia arvoja yms. Toinen kysely esittää sukupuolikoodin jakaumia ja paljastaa myös, paljonko on puuttuvia tai esim. sukupuoli- koodin 9 tai blankko esiintymiä.

Yhdistämisellä tarkoitetaan eri lähteistä tule- vien, samaa tarkoittavien tietojen vertailua ja integrointia. Tyypillinen esimerkki on asiakastie- tojen yhdistäminen. Monitoroinnilla ymmärretään tietojen laadun jatkuvaa valvontaa.

Menetelmien toteuttamiseksi voidaan käyt- tää erityisiä tietojen profilointiin, yhdistämiseen, monitorointiin yms. toimintoihin erikoistuneita ohjelmistoja. Esimerkkejä ovat Trillium, DataFlux ja IBM Quality Stage. Lisäksi monet ETL-välineet sisältävät laatuun liittyviä osia. Toistaiseksi näiden ohjelmistojen käyttö Suomessa on ollut vähäistä. Saattaa kuitenkin käydä kuten ETL-välineiden kohdalla: käyttö lisääntyy kunnes perinteistä ohjelmointia ei enää edes harkita.

Organisointi

Asiat eivät synny ilman tekijöitä. Tiedonhallin- nan alueella on monenlaisia töitä ja tarvitaan eri- laisia organisaatioyksiköitä ja työtehtäviä.

Yksi ajankohtainen alue on Data Governance eli hyvä tietojenhallintatapa. Tälle alueelle kuuluvat laadun lisäksi integrointi-, Master Data, Metadata ja tietovarastointiasiat. Data Governancea varten on monissa edistyksellisissä taloissa jo perustettu oma organisaatioyksikkö, joka toimii yli organi- saatorajojen.

Yhteenveto

Tietojen laatu liittyy yleiseen tietoresurssin arvostukseen. Huono laatu on usein piilevää ja paljastuu tietovarastointi- ja konversiohan- keissa. Olisi syytä laatia yrityksen laatustrategia, jossa määritellään mm. tavoitteet tietojen laa- dulle. Laadun parantamiseksi on erilaisia menetel- miä kuten profilointi ja monitorointi. Markkinoilla on myös hyviä ohjelmistoja automatisoimaan eri menetelmiä, niihin kannattaa tutustua. Osana tie- donhallinnan organisointia myös laatuasiat tarvit- sevat tekijänsä.



Kirjoittaja: Ari Hovi
Ari Hovi Oy

