

SYSTEMITYÖ

Systeemityöyhdistys SYTYKE ry:n jäsenlehti N:o 4/2005



Teemana: Tiedon ja tietämyksen
hallinta



Information Solutions for a Global Age

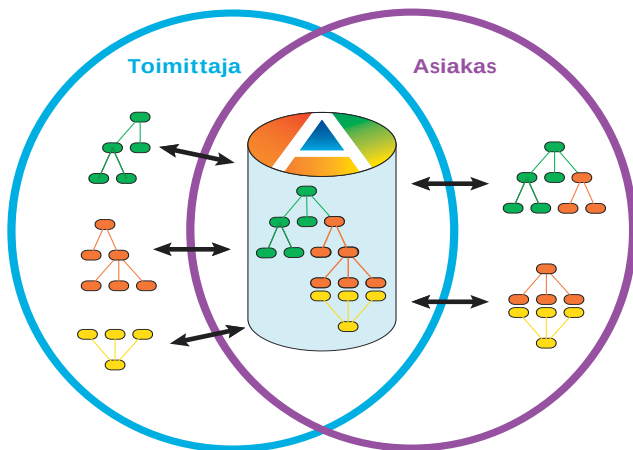
Asiantuntijapalveluja ja tietojärjestelmäratkaisuja tuotetiedonhallintaan Suomessa vuodesta 1995.

Asiantuntijapalvelut:

- Projekti- ja hankehallinta, vaatimustenhallinta, laadun varmistus
- Mallintaminen (tuotetieto, prosessit, järjestelmäarkkitehtuuri)
- Kansainvälisten tuotetietostandardien koulutus (mm. STEP, IFC, PLCS)
- Tietojärjestelmien määrittely ja toteuttaminen

Tiedonhallintaratkaisut:

- Share-A-space (www.share-a-space.com)



Jaa tieto yhdestä paikasta

Share-A-space auttaa hallitsemaan tuotetietoa riippumatta siitä, missä se tuotetaan tai missä sitä käytetään. Share-A-space on tietokeskus, joka integroi eri lähteistä ja tietojärjestelmistä tulevan tiedon standardiformaatissa.

Volvo Penta, Land Systems Hägglunds, FMV ja Saab Bofors Dynamics ovat yrityksiä, jotka käyttävät Share-A-spacea tehostamaan jokapäiväistä toimintaansa. Näillä yrityksillä on useita toimittajia, asiakkaita ja pitkäikäisiä tuotteita. Tehokas tiedon yhteiskäyttö tuotteen koko elinkaaren ajan parantaa yrityksen tuottavuutta ja avaa uusia liiketoimintamahdollisuuksia.

-Toolbox -luokkakirjastot

Toolbox luokkakirjastot tarjoavat ratkaisun tietojärjestelmien (CAD, PDM, ERP) väliseen tiedonvaihtoon standarditietomalleihin STEP, IFC, PLCS perustuen.

Eurostep Finland, Karvaamokuja 3, 00380 Helsinki
tel: 09-3487 2030 fax 09-3487 2031
email: info_fi@eurostep.com www: www.eurostep.com

Julkaisija

Systeemityöyhdistys Sytyke ry
Puhelinvastaus- ja sihteeripalvelu VT Oy
Susanna Koskinen
Henrikintie 7 A, 00370 Helsinki
p. 09 5607 5363
f. 09 5607 5365
sytyke@hennax.fi

Päätoimittaja

Lauri Laitinen, Nokia
lauri.laitinen@nokia.com
puh. 050 483 6551 (NRC)

Toimitussihteeri

T:mi M. Pyhäjärvi
Marika Pyhäjärvi
PL 119, 00701 Helsinki
p. 044 333 2558
marika.pyhajarvi@gmail.com

Toimituskunta 4/2005

Hanna Horppila
Tomi Kauppinen
Jouni Lähteenmäki
Minna Oksanen
Jukka Ponsi
Kari Sipilä

Lisätietoja lehdestä

<http://www.sytyke.org/lehti/>

Tilaukset

Systeemityölehti sisältyy
yhdistyksen Tietotekniikan liiton
suositusten mukaiseen
yhdistyksen jäsenmaksuun.
Vuositilaus 20 €
Irtotilaukset 5 €

Hyvissä ajoin ennen painatusta tehty vä-
hintään 50 kappaleen lisätilaus 2 €/kpl.
Tilaukset yhdistyksen
toimistosta.

Kansikuva

Jeffrey Garrett
The Tree of Porphyry, one of the earliest
examples of a concept hierarchy. From
the ceiling fresco at Schussenried, Ger-
many, by Franz Georg Herrmann (1757),
photograph by Jeffrey Garrett (North-
western University Library, October
2000).

Painopaikka

T-Print
Ahokaari 1-3
05460 Hyvinkää
Puh. (019) 475 8500

Painos: 2500 kpl

ISSN 1237-0525
12 vuosikerta, no. 4



PÄÄKIRJOITUS

- 4 Hanna Horppila: Uusi osaamisyhteisö Suomessa

TEEMA-ARTIKKELIT

- 5 Aija Palomäki: Informaatioarkkitehtuurityö Nokialla
9 Anna Aaltonen ja Jouni Lähteenmäki: Yritystietomalli – haaveita ja haasteita
13 Henry Runn: Tietovaraston sovellusarkkitehtuuri muutoshallinnan näkökulmasta
16 Minna Oksanen: Sirpaletiedosta kokonaisuuteen
21 Päivi Happonen: Tiedon elinkaari hallintaan
23 Markus Merenmies: Arkistolaitoksen sähköinen arkistointi- ja tietopalvelujärjestelmä
27 Jonne Kulluvaara: Musiikin metadata Yleisradion äänilevystössä
29 Veli-Matti Kiviranta: Paikkatieto – avain tiedon kokonaisvaltaiseen hyödyntämiseen
32 Tomi Kauppinen, Tuukka Ruotsalo ja Mirva Salminen: Tietämyksen mallintaminen semanttisessa webissä
35 Leyla Akgez: Tiedonhallinnan ontologiaa tekemässä
38 Pekka Valta: Ensikokemuksia ontologioista
41 Viljo Martikainen: Tietämys ihmisen mentaalisenä ilmentymänä
45 Paula Miinalainen: Hiljainen tieto yrityksen voimavarana

YHDISTYS

- 48 Tarja Raussi: Arkkitehtuurit – taidetta vai tekniikkaa

Ilmoitushinnat

Koko	Mustavalko	4-väri
Takakansi A4	-	1200 €
Sisäkannet A4	-	1000 €
Sisäsivut 1/1	400 €	800 €
Sisäsivut 1/2	200 €	600 €
Sisäsivut 1/4	100 €	-

Arvonlisävero 0 %

Vakiopaikan vähintään vuodeksi varanneille 20 % alennus.

Uusi osaamisyhteisö Suomessa

Hanna Horppila, Dama Finland

Tämän lehden teema, tiedon ja tietämyksen hallinta, on erittäin moniulotteinen. Se kuvastaa hyvin sitä aihepiiriä, josta kiinnostuneet ihmiset perustivat vuosi sitten DAMA Finland -nimisen osaamisyhteisön. Miksi miettiä yksin alan uusia suuntauksia ja menetelmiä, kun voi olla yhteydessä saman alan ammattilaisiin riippumattoman verkoston kautta?

Osaamisyhteisön idea on jakaa ja etsiä yhdessä tietoa hyvistä toimintatavoista, käytännöistä ja työkaluista. Se kokoaa yhteen ihmisiä, joilla on samankaltaiset mielenkiinnon kohteet. Jokaisella on ainutkertainen työ- ja kokemusprofiilinsa. Kun olemme aktiivisia antamisessa, saamme monin verroin takaisin. Osaamisyhteisö on virtuaalinen opinahjo sanan parhaassa merkityksessä.

DAMA Finlandin perustamista valmisteli puolen vuoden ajan toimikunta, joka toi vuosi sitten lokakuussa perusteellisen esityksensä käytännön asioiden järjestämisestä ja vaihtoehtoisista malleista osaamisyhteisön järjestäytymisen ja käytännön asioiden hoitamisen osalta Nokian tiloissa järjestettyyn kokoukseen, jonne oli kutsuttu sähköpostilla suuri joukko tiedonhallinnan ammattilaisia. Tuo kokous päätti DAMA Finlandin perustamisesta ja sen organisoimisesta SYTYKE ry:n osaamisyhteisöksi sekä kansainvälisen DAMA Internationalin jaostoksi.

Ensimmäinen toimintavuosi on ollut vauhdikas. Kohta organisointumisen jälkeen ryhdyimme järjestämään jäseniltoja asiantuntijaluentoja ryydittämällä. Kirjoitimme artikkeleita myös ITviikkoon.

Yhteisön jäsenlistalla on tällä hetkellä noin 200 henkilöä.

Toiminnan tiimoilta tuli ilmi, että tiedon ja tietämyksen hallinnalla tarkoitetaan eri organisaatioissa hieman eri asioita. Systemaattisina ihmisinä damalaiset ryhtyivät jäsentämään tietojenhallinnan käsitteistöä työseminaarien avulla. Tästä ontologiatyöstä kerrotaan tämän lehden artikkeleissa ja siihen pääsee mukaan seuraamalla DAMA Finlandin sivuilta www.damafinland.org löytyviä ohjeita. Tervetuloa mukaan muuhunkin DAMA Finlandin toimintaan.

Näiden artikkelien kirjoittajat lähestyvät lehden teemaa useista eri näkökulmista – lähtien datan hallinnasta ja jalostamisesta aina tietämykseen ihmisen mentaalisenä ilmentymänä. Tarjolla on siis sekä käytäntöä että teoriaa. Kirjoittajat ovat löytäneet aktiivisten damalaisten verkostojen kautta. He ovat paneutuneet asiaansa. Olkaa hyvät ja nauttikaa lukemastanne.

Hanna Horppila työskentelee tietoasiantuntijana Tieliikelaitoksen konsultoinnissa.

Hanna.Horppila@tieliikelaitos.fi

Informaatioarkkitehtuurityö Nokialla

Aija Palomäki, Nokia



Kerron tässä artikkelissa nokialaisesta informaatioarkkitehtuurista, ja arkkitehtuurivisiosta yleisemminkin. Pyrin jopa määrittelemään sanan arkkitehtuuri! Kuvaan myös lyhyesti keskeisiä arkkitehtuuriajattelun jalkauttamisen menetelmiä.

Nokialaisen informaatioarkkitehtuurin tärkein komponentti on yritystason tietomalli. Koska ison talon globaalin liiketoiminnan tietosisällön yksityiskohtainen mallintaminen on haastavaa ja aikaaviepää, tietomalli on korkealla abstraktiotasolla, mutta sisältää joukon uudelleenkäytettäviä mallirakenteita (data design patterns) käytännön mallinnushaasteiden ratkomiseen.

Mitä arkkitehtuuri on?

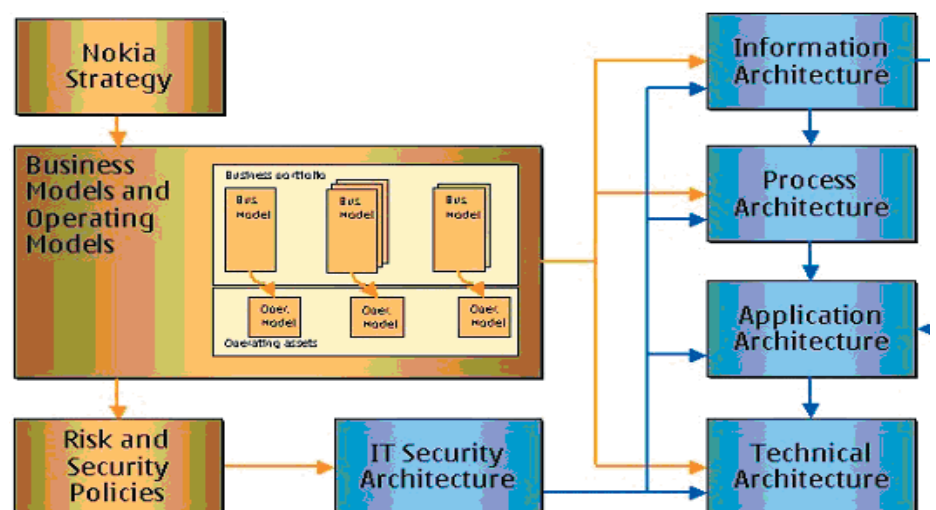
Jo termodynamiikan lakienkin mukaan mikä tahansa järjestelmä alkaa hajaantua kaaokseksi, jollei järjestelmään tuoda ulkopuolista energiaa. Näemme arkkitehtuurin juuri tuona järjestystä luovana energiana – järjestys luodaan rakenteilla, joita myös arkkitehtuureiksi kutsutaan.

Sanalla arkkitehtuuri tarkoitetaan aina isoa kuvaa, helikopterinäköä. Näkymä voi liittyä informaatioon, prosesseihin, sovelluksiin, tietoturvaan tai teknisiin alustoihin.

Arkkitehtuurivisio

Nokia Business Infrastructures –organisaatio sisältää mm. nokialaisen tietojärjestelmätyön ja –tuen.

Nokia BI on lähtenyt toteuttamaan arkkitehtuurivisiota kahdesta tärkeästä syystä: tehostakseen oman toimintansa tuottavuutta ja



Kuva 1. BI arkkitehtuurin dimensiot.

nopeutta suhteessa liiketoiminnan tarpeisiin.

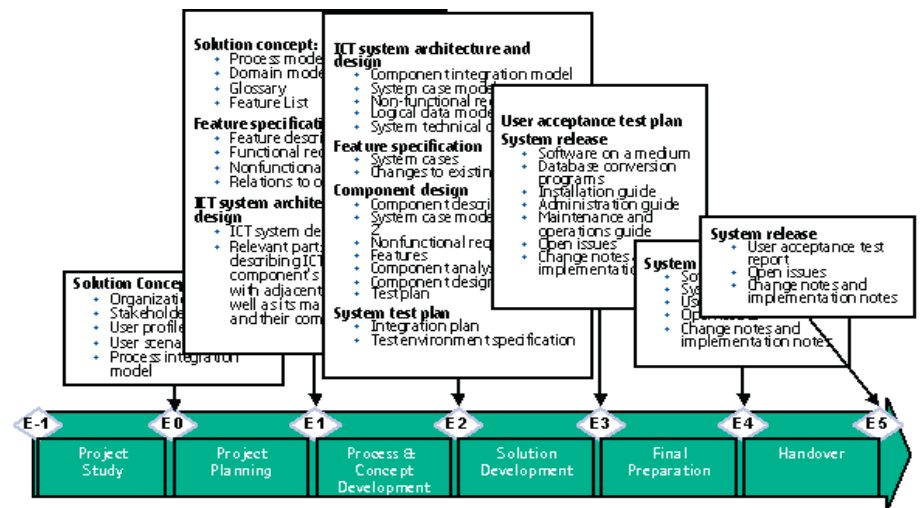
Nokia BI arkkitehtuuri perustuu Nokia-tason strategiseen arkkitehtuuriin, ja se koostuu viidestä eri dimensiosta (ks. kuva 1).

Käytännössä BI-tason arkkitehtuurityö tuottaa ohjeita ja standardeja, jotka ohjaavat toteuttamaan liiketoimintamallista johdettuja ns. ei-toiminnallisia vaatimuksia. Ei-toiminnalliset vaatimukset liittyvät tyypillisesti tietojärjestelmien

- näkyvyyteen ja yhdistettävyyteen
- optimointiin
- suorituskyykyyn ja kapasiteettiin
- saatavuuteen ja jatkuvuuteen
- käytettävyyteen ja operoitavuuteen
- joustavuuteen, laajennettavuuteen, ylläpidettävyyteen ja
- mobiilisuuteen.

Arkkitehtuuridimensioista lyhyesti:

- Informaatioarkkitehtuuri – tämän artikkelin pääaihe joten siitä lisää myöhemmin.
- Prosessiarkkitehtuuri – helikopterinäkymä keskeisiin liiketoimintaprosesseihin ja niiden keskenään vaihtamaan tietoon
- Applikaatioarkkitehtuuri – helikopterinäkymä applikaatioiden roolijakoon sekä niiden keskenään vaihtamiin tietoihin
- Tietoturva-arkkitehtuuri – BI-tason tietoturvastrategiat ja -standardit



Kuva 2. NOCOP etapit ja UNA välituotteet

- Tekninen arkkitehtuuri – BI-tason teknologian ja teknologiavalintojen hallinta

Jalkauttaminen

Arkkitehtuuridimensioiden ohjeistot ja standardit joko toteutuvat tai jäävät toteutumatta käytännön tietojärjestelmäkehityksessä, projektityössä. Näiden projektien aikataulut ja tavoitteet ovat usein varsin tiukat, ja siten laatu- ja arkkitehtuuriajatteluun käytettävissä oleva aika on rajallinen.

Keskeiseksi onnistumistekijäksi nousee siis projektien tehokas menetelmäohjeistus. BI:ssa menestymistä jaetaan kahteen eri kerrokseen:

- Projektinhallintamenetelmistö, NOCOP (management)
- Tietojärjestelmätyö, UNA (engineering)

NOCOP määrittelee, missä projektin etapissa minkäkin UNAn mukaisen tuotoksen tulee olla valmis ja mitkä roolit työhön osallistuvat. Lisäksi BI-tasoinen arkkitehtuurielin valvoo että edel-

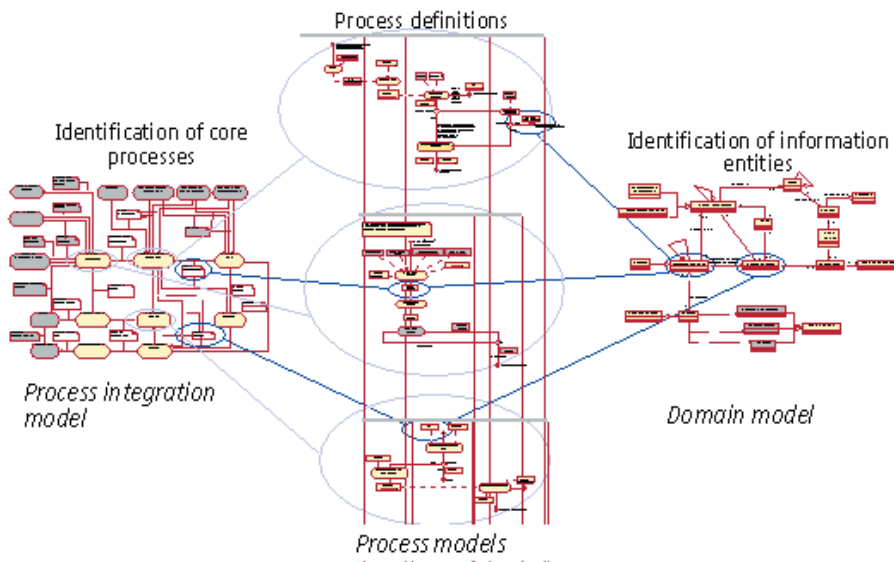
lytetyt tuotokset ovat syntyneet ja ovat laadultaan sopivat.

UNA määrittelee puolestaan millaisia tuotoksia tietojärjestelmätyössä tarvitaan ja mitkä ovat niiden tuottamiseen tarvittavat syötteet sekä tuotosten väliset riippuvuussuhteet. UNA sisältää myös työohjeet ja mallit. Kuvassa 2 näkyvät NOCOPin etapit suhteessa UNA välituotteisiin.

UNA kehitettiin BI-tason arkkitehtuurifunktion johdolla vuosina 2002-2003, alun perin RUP- (Rational Unified Process) pohjalta, mutta niin, että mukaan tuotiin voimakas arkkitehtuurinäkemys. RUPin tapaan UNAn on mallipohjainen – dokumentit generoidaan keskitetystä mallirepositoriosta käsin.

UNA hyödyntää UML (Unified Modeling Language) -notaatiota kaikessa mallinnuksessa, mukaan lukien prosessi- ja tietomallit.

UNasta tehtiin RUPista poiketen minimaalinen sekä voimakkaasti integroitu siten, että esim. proses-



Kuva 3. UNAn keskenään integroituja välituotteita.

sikuvauksissa (process model) identifioituille tiedoille löytyy aina tietomalli (domain model), ja prosessikuvausten automatisoidulle aktiviteeteille käyttötapausmalli (use case model), ks. kuva 3.

Keskeisin muutos UNAssa verrattuna BI:ssa aiemmin käytössä olleisiin kehitysmenetelmiin on se, että vaatimukset ja tulevan järjestelmän arkkitehtuuri tulee määritellä jo varhaisessa, projektin suunnittelu-, -vaiheessa. Näin saadaan aikaan mm. realistiset kustannusarviot.

Informaatioarkkitehtuuri

Nokialaisen informaatioarkkitehtuurin tärkein komponentti on yritystason tietomalli, NEIM (Nokia Enterprise Information Model).

NEIM määrittelee keskitetysti

- käytetyn terminologian (nimet ja määritelmät) ja
- suositellut tietorakenteet.

NEIMin kuvauskieli on UML. NEIM sisältää metamallin, käsitteellisiä malleja viidellä eri

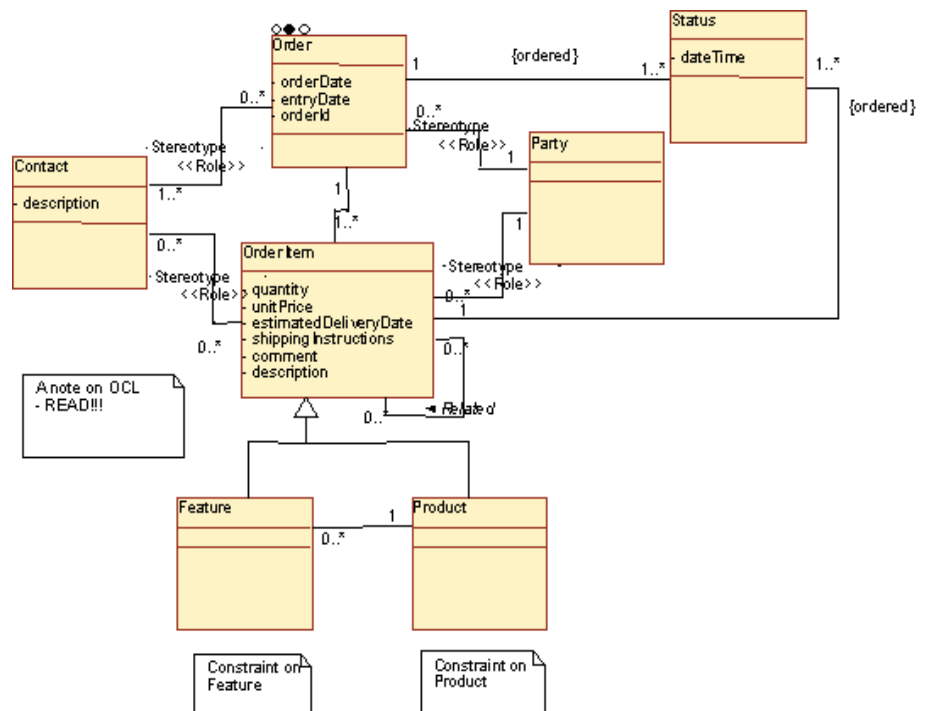
osa-alueella - osapuolet (party), tapahtumat (transaction), tuotteet (product), työ (work effort) ja talous (financial management) – sekä loogisia tietomalleja. Kuvassa 4 esimerkki NEIM-mallista (tilaus).

NEIM- kehitystyössä on nojaututtu valmiiseen materiaaliin niin paljon kuin se on ollut käytännössä mahdollista. Nokia BI on mm.

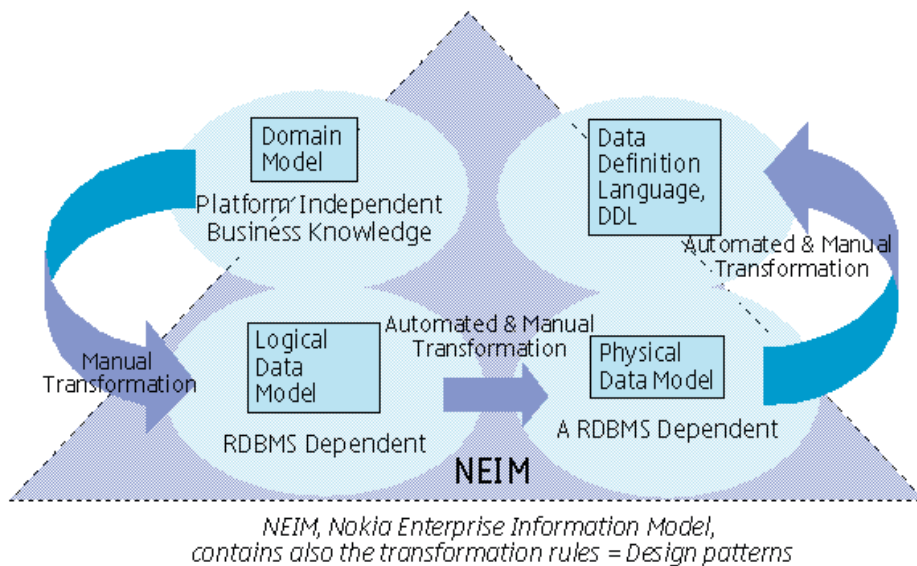
lisensoinut UDM (Universal Data Models) mallirepositorion, validoinut sen sisältöä Nokian liiketoimintamalleja vasten, ja yhdistänyt näin syntyneet räätälöidyt mallit NEIMiin.

NEIM noudattaa myös OMG:n (Object Management Group) MDA –ajattelua (Model Driven Architectures) siten, että käsitetasolla kuvataan vain keskeisin liiketoimintasemantiikka, ja muunnosta käsitteelliseltä tasolta loogiselle ohjaavat valmiiksi määritellyt muotit (patterns) – samoin kuin muunnosta loogisesta tietomallista fyysiseenkin ja edellen tietokannan kuvauskielen. Tällä hetkellä MDA on pitkälle vain ajattelumalli – tulevaisuudessa mallinnustasojen väliset muunnokset voitaneen myös automatisoida.

Muotit ovat erittäin keskeinen osa NEIMiä – tarjoamalla joukon standardeja ratkaisumalleja tieto-



Kuva 4. Tilauksen tietomalli



Kuva 5. NEIM ja MDA

mallinnushaasteisiin, yhtenäistämme yrityksen tietomallinnusta niissäkin tapauksissa, että projektityössä tehty mallinnus tapahtuu sellaisilla liiketoiminnan alueilla, joihin NEIMissä ei vielä ole olemassa valmista tietomallia.

NEIMiä kehitetään julkaisuina, ja tällä hetkellä on menossa julkaisu numero 3. NEIM-julkaisut hyväksytään katselmointiprosessin kautta, ja julkaistaan ns. BI Reuse -mallirepositoriossa, jolloin ne ovat kaikkien BI projektien käytössä. UNA-ohjeistojen ja arkkitehtuurin katselmointiprosessin kautta varmistetaan, että projekteissa tuotettavat tietomallit ovat yhteensopivat NEIM-määrittelyjen kanssa.

NEIMin kehitystä ohjaavat liiketoimintaprosessit ja niiden implikoimat tietotarpeet. Kaikki NEIMiin tulevat mallit pilotoidaan käytännössä, jolloin saadaan palautetta niiden toimivuudesta. Tällä tavoin yritystason tietomalli kehittyy koko ajan täydellisemmäksi, koska siihen kumu-

loituu tietämystä tietovaatimuksesta eri puolilta organisaatiota. Tässä lähestymistavassa joudutaan kuitenkin hyväksymään se, että yritystason tietomalli on lähtökohtaisesti epätäydellinen, ainoastaan ”riittävän hyvä”, ja se kehittyy jatkuvasti käytännön kokemuksen kautta mm. niin että tietomalli alkuvaiheessa sisältää vain keskeisen rakenteet ja yksityiskohdat täydennetään seuraavissa julkaisuissa. Näin vältetään ns. ”analysis paralysis”-syndrooma.

Johtopäätöksiä

Kehitettyämme nyt parin vuoden ajan keskitettyä yritystason tietomallia, olemme oppineet mm. seuraavaa:

- Tietomallin kehitys ei saa olla akateeminen harjoitus vaan sen on vastattava selkeään liiketoimintatarpeeseen
- Tietomalli kannattaa aina ostaa, mikäli mahdollista, tyhjästä alkava kehitys on aikaavieppää ja kallista
- Yritystason määritelmät vaativat laajan hyväksynnän –

vaarana kuitenkin aina ”analysis paralysis”

- Menestystarinat ovat tärkeitä – kehitettyä arkkitehtuuria tulee voida soveltaa nopeasti johonkin todelliseen liiketoimintaongelmaan
- Vaikka arkkitehtuurityön toteutumista hankkeissa tuleekin valvoa, arkkitehtuuri ei saa olla poliisifunktio vaan ennemminkin mahdollistaja – arkkitehtien tulee mennä hankkeisiin varhain, mieluummin uuden etapin alussa kuin lopussa, tehdä odotukset selviksi ja mahdollistaa tavoitteisiin pääsy auttamalla itse aktiivisesti.

Lähteet

MDA - Model Driven Architectures, ks. <http://www.omg.org/mda/>

OMG – Object Management Group - <http://www.omg.org/>

RUP – Rational Unified Process - <http://www-306.ibm.com/software/awdtools/rup/>

UDM - Universal Data Models - <http://www.univdata.com>

UML - Unified Modeling Language - <http://www.uml.org/>

Artikkelin kirjoittaja Aija Palomäki työskentelee informaatioarkkitehtina Nokiassa, vastuualueellaan tällä hetkellä Nokian tietovaraston tietomallien kehittäminen. aija.palomaki@nokia.com

Yritystietomalli – haaveita ja haasteita

Anna Aaltonen, Gapgemini & Jouni Lähteenmäki, Op-keskus

Artikkelissa pohditaan yritystietomallille yleisesti asetettuja tavoitteita ja sitä voivatko nämä tavoitteet toteutua yhdessä ja samassa yritystietomallissa. Eri käyttäjäryhmillä on omat tarpeensa ja oma näkökulmansa siihen, miten asiaa pitäisi edistää. Artikkelin kuvaa käytännön asioita, joihin yritystietomallia suunnittelevat törmäävät.

Onko selvää mitä yritystietomallilla tarkoitetaan?

Yritystietomallia pidetään yleisesti hyödyllisenä ja kannatettavana asiana. Jos joku ehdottaa palaverissa, että ”meillä pitäisi olla yritystietomalli”, asiasta ollaan helposti yhtä mieltä. Erityisesti, jos toteamus on sen verran yleinen, ettei tehtäviä ja vastuita yritystietomallin toteuttamiseksi aleta jakaa. Kullakin osallistujalla on oma käsityksensä siitä, mitä yritystietomalli tarkoittaa ja mitä hyötyä siitä olisi. Yritystietomallilla ajatellaan saatavan aikaiseksi yhdeksän hyvää ja kahdeksan kaunista. Jotkut ovat aikoinaan osallistuneetkin yritystietomalli-projekteihin – tosin nämä henkilöt suhtautuvat asiaan huomattavasti muita skeptisemmin.

Osa syistä, miksi yritystietomalli-projekteissa ei päästä tavoitteisiin on se, etteivät tavoitteet ole riittävän selkeät. Vaikka tavoitteita kirjoitettaisiin projektisuunnitelmaan, jäävät ne helposti maailmoja syleileviksi.

Tarve yritystietomallille havaitaan useimmiten kantapään kautta. Ollaan integroimassa järjestelmiä yhteen ja huomataan, että rajapinnoissa joudutaan tekemään melkoinen määrä konversioita tiedoille. Tiedot periaatteessa löytyvät, mutta eri muodossa. Asian selvittäminen edellyttää melkoista tietomallien läpikäyntiä. Pahimmillaan konversiosäännöt ovat vaikeita tai mahdottomia tehdä. Periaatteessa molemmissa järjestelmissä on samoja tietoja,

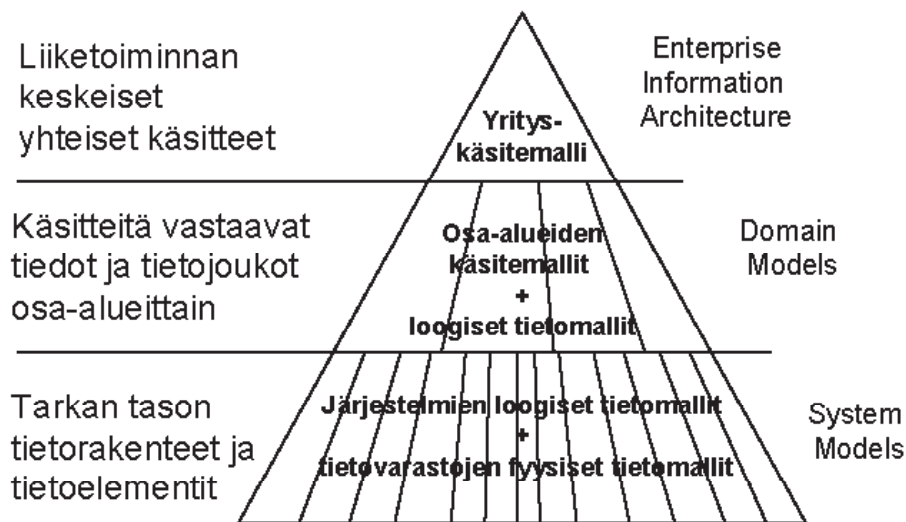
kuten se kuuluisa ”asiakas”. Tällöin on luonnollista kysyä, eikö se voisi olla se sama ”asiakas”? Päätetään siis, että jatkossa talletetaan yritystietomalliin tämä yhteinen käsite ja maailma pelastuu.

Mitä eri kehittäjäryhmät odottavat yritystietomallilta?

Käyttötarve määrää pitkälti sen, mistä päästä, mihin tarkoitukseen ja millä tarkkuustasolla yritystietomallia aletaan rakentaa. Tarpeet vaihtelevat käyttäjittäin, yksi haluaa yhtä ja toinen toista. Kaikki asiat ovat sinänsä kannatettavia, mutta mahdottomia toteuttaa yhtäaikaan. Erilaisia tarpeita, joihin ratkaisuksi tarjotaan yritystietomallia ovat:

- kaivataan ”yhteistä kieltä”, esimerkiksi liiketoiminnan edustajien ja teknisten toteuttajien välille tai eri liiketoimintalueiden välille. Tähän ratkaisu on yhtenäinen sanasto. Sanasto toteutetaan perinteisesti hilitömmän pitkänä excel-tiluk-kona, jonka ylläpito tyssä omaan mahdottomuuteensa melko nopeasti. Sanaston tallennuspaikan lisäksi ongelma on eri nimiavaruuksien hallinta. Eri sidosryhmät käyttävät samaa termiä eri asioista, esimerkkinä tästä esimerkiksi sana ”palvelu”. Sana on sidosryhmän sisällä niin vakiintunut termi, että sen korvaaminen toisella olisi keinotekoisia ja vaatisi melkoista uudelleenopetteluä kaikilta osapuolilta. Yrityksissä ei liene järjestetty

Yrityksen tietojen kuvaustasot



Kaavio 1. Yrityksen tietojen kuvaustasot

- koskaan sanakokeita termistöstä, kuten koulussa kielistä. Jos sanastossa pystytään huomioimaan nimiavaruudet, näiltä sanakokeilta sitäpaitsi välttään.
- määrittelijät tekevät loogisia käsitelmalleja ja kaipaavat lähinnä uudelleenkäytettävää materiaalia. Tuntuuhan turhautavalta keskustella kerran toisensa jälkeen käsitteistä, jotka on varmasti määriteltä jo moneen kertaan. Liittymien läpikäynnissä tämä korostuu erityisesti, voisihan määrittely keskitä aidosti uusiin asioihin eikä vanhojen kaiveluun. Tässä tulee ongelmaksi se, että vaikka asia on aikaisemmin kuvattu on se kuvattu hieman eri näkökulmasta ja eri tarkoitukseen. Kun kuvaukset vielä ovat pitkiä tekstiasiakirjoja, ei kuvauksen yhden osan poimiminen uudelleenkäytettäväksi onnistu kovin helposti muutenkaan.
- tietokantojen ja rajapintojen suunnittelijat taas kaipaavat melko tarkankin tason määrittelyjä, jotka voisi suoraan ottaa pohjaksi, kun generoi taulurakenteita. Yritystietomallin käyttö suoraan toteutuksessa toki on varma tapa varmistaa, että yhteisiä käsitteitä todellakin tulee käytettyä. Tarkan tason tietomalli on kuitenkin työläämpi tehdä ja osa asioista voi olla teknologiariippuvaisia. Näitä tarpeita vastaavat yrityksen tietojen kuvaustasot, jotka on esitetty seuraavassa kaaviossa 1.

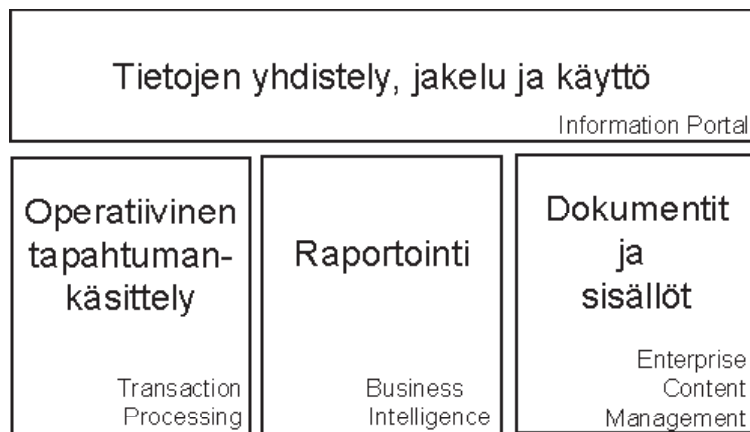
Miksi yritystietomallin jäsentäminen on niin vaikeaa?

Yrityksen kullakin liiketoimintalueella on tyypillisesti erilaiset tulkinnat keskeisistä käsitteistä ja erilaiset riippuvuudet niiden välillä. Esimerkiksi finanssisektorin yrityksessä käsite "asiakas" tarkoittaa varsin erilaista asiaa sijoituspalveluiden puolella kuin perintäpalveluissa. Ihmiset pystyvät useimmiten intuitiivisesti ymmärtämään

toisiaan käsite-eroista huolimatta, mutta formaali käsitelmallinnus ei voi onnistua jos väkisin samaksi käsitteeksi yritetään yhdistää aivan eri asiat, joista vaan satutaan käyttämään samaa termiä.

Lisäksi eri käyttäjäryhmillä voi olla erilaisia termejä samoille asioille ja nämä käyttäjäryhmäkohtaiset käsitteet voivat ryhmittyä aivan eri tavoin. Esimerkiksi markkinointiosastolla voidaan tuotteesta haluta puhua asiakkaille "palveluna". Pankin tuottama "käyttötili" kaikkine siihen liittyvine toimintoineen ja käyttökanavineen voidaankin asiakkaille markkinoida "päivittäisten raha-asoiden palveluna", kun taas asiakkaalle palvelun myyvä toimihenkilö tai verkkosivusto kertoo samasta tuotteesta sopimuskäsittein. Edelleen tuotteita ja tuotepaketteja kehittäville ihmisillä on tuotteille omanlaisensa tyypittely ja tuotteen osien ryhmittely, joka lähtee palvelun tuoteistajan ja palvelun pyörittämisen näkökulmasta. Mikäli käsitteiden mallintaja ei tunnista näitä näkökulmaeroja, ei ole ihme jos yrityksen tietojen mallinnus tuntuu hankalalta.

Tietojärjestelmiä kehitettäessä hyödylliset käsitelmallit taas riippuvat toteutettavan järjestelmän tyypistä. Tiedot mallinnetaan varsin eri tavoin riippuen siitä, onko kyseessä esimerkiksi asiakasportaalin sisältöjen ja palvelujen navigaatorakenteen vaiko vakuutusmaksujen kirjanpidon toteutus. Erityyppiset tietojärjestelmät käsittelevät eri tasoista tietoa: datan, informaation ja tietämyksen mallinnuksessa tunnis-



Kaavio 2. Tietojen tyypittely niiden käyttötarkoituksen mukaan

tetaan erilaisia käsitteitä. Lisäksi täsmälleen sama tietoelementti voi toisen järjestelmän kannalta olla dataa, toisen kannalta metadataa.

Kaaviossa 2. on esitetty mallituksen kannalta luonteeltaan erityyppiset tietojenkäsittelytarpeet artikkelin kirjoittajien kokemuksen mukaan.

Yritystietomallin käsitteiden suhteuttaminen yksittäisen tietojärjestelmän tarkan tason tietorakenteisiin voi näkökulmaeroista johtuen olla vaikeaa. Vastaavasti yhden tietojärjestelmän loogisen käsittemallin yleistys yritystietomalliksi voi olla vaarallisen yksiselimäistä. Yrityksen keskitetty DW toki sisältää jonkinlaisen kokonaiskuvan yrityksen tiedoista, mutta on muistettava että raportointijärjestelmät on tehty vain tietynlaisia tarpeita. Muunlaisten tietojärjestelmien tekijöille tällaisesta yritystietomallista ei välttämättä ole juurikaan hyötyä.

Näkökulmat yritystietomalliin ja osa-alueiden tunnistaminen

Yrityksen eri henkilöillä on siis tyypillisesti hyvin erilainen näkemys yrityksestä ja sen tiedoista

kokonaisuutena. Se on luonnollinen seuraus toimenkuvan mukaisesta paikasta organisaatiossa ja sen mukaisesta näkökulmasta. Mikäli yhteen malliin yritetään kaikki näkökulmat yhdistää, lopputuloksena on helposti niin yleinen kuvaus tai niin sekalainen kompromissien kokoelma ettei siitä ole kenellekään käytännön hyötyä.

Kaaviossa 3 on pyritty jäsentämään yritystietomallin kannalta keskeisiä näkökulmaeroja. Kaavion ideana on havainnollistaa erilaisia

Liiketoiminta-näkökulma

Toimialakohtaiset käsitteet
 • jokaisella yrityksen liiketoiminta-alueella erilaiset käsitteet ja niiden suhteet

Käyttäjä-näkökulma

Käyttäjryhmäkohtaiset käsitteet
 • asiakasryhmät
 • toimihenkilöryhmät
 • kehittäjäryhmät
 • eritasoiset johtajat

Tekninen näkökulma

Erityyppiset järjestelmät
 • informaatioportaali
 • tapahtumankäsittely
 • raportointi
 • dokumentit ja sisällöt

Kaavio 3. Erilaiset näkökulmat yrityksen tietoihin

tarpeita yritystietomallille ja auttaa hahmottamaan yritystietomallin monitahoista luonnetta. Tällaisen analyysin perusteella voidaan tunnistaa ja valita ne osa-alueet, jolle oma yritystietomalli halutaan kohdistaa. Osa-alueiden ja kuvausnäkökulmien valinnassa on mietittävä kenelle ja mitä hyötyä yritystietomallilla ollaan tavoittelemassa.

Miten eri osa-alueet suhtautuvat toisiinsa?

Vaikka edettäisiin vain valituilla osa-alueilla – koko maailmaa kun ei pysty laittamaan kuntoon kerralla – olisi löydettävä tasapaino. Yhtä näkökulmaa ei voi viedä eteenpäin kovinkaan pitkälle yksinään. Sitäpaitsi kaikki näkökulmat täydentävät toisiaan, kannattaa aloittaa yhdeltä liiketoiminta-alueelta ja huomioida parin käyttäjäryhmän tarpeet. Yleensä toteutuslupustakin on joku tietty. Jostain kun on aloitettava, niin

väkisinkin tulee valittua asioita kustakin näkökulmasta.

Jos yritystietomalli koostetaan tällä tavalla erilaisista palasista, onkin hyvä kysymys, onko meillä yksi vai useampi yritystietomalli? Jos on useampi yritystietomalli, niin onko se edes yritystietomalli vai onko se perinteinen järjestelmäkohtainen tietomalli, johon olemme tottuneet tähän mennessä? Hyviä kysymyksiä, mutta melko periaatteellista pohdiskelua.

Miten yritystietomallin kanssa kannattaisi edetä?

Tässä artikkelissa on mainittu monia asioita, joita yritystietomalli helpottaisi. Kunkin on mietittävä omassa organisaatiossaan, mitkä ovat painopisteet. Tehdäänkö ohjelmistot itse vai käytetäänkö valmispaketteja? Jos valmispaketin mukana tulee tietomalli, kannattanee tyytyä pelkän loogisen tason käsitteiden määrittelyyn itse. Ja vaikka päätyisi laatimaan kattavamman yritystietomallin, kannattaa ehdottomasti etsiä valmiita käsitteiden jäsennyksiä työn pohjaksi eikä pyrkiä keksimään kaikkea itse.

Tekeepä yritystietomallia mihin tarkoitukseen tahansa, tarvitaan siinä välinetukea. Yritysmallin on sovelluttava monen ihmisen luettavaksi ja päivitettäväksi; pitäisihän sen olla "elävä", laajalle levinnyt työväline. Pelkkien ohjeiden varassa suoritettavat manuaalitehtävät (tiedostoihin tai taulukoihin tehtävät päivitykset) eivät takaa riittävän yhtenäistä laatua. Jos tehtävä on liian työläs se jää ennemmin tai myöhemmin tekemättä. Kun

yritystietomallin rapautuminen on kerran alkanut, vauhti vain kiihtyy. Motivaatio ylläpitää mallia, jonka tietää olevan jo vanhentunut, on yleensä olematon.

Yritystietomallin kehittäminen kannattaa liittää yrityksen kokonaisarkkitehtuurin kehittämiseen (Enterprise Architecture). Yhtenäiset tiedot ovat se avain, joka avaa tien yrityksen liiketoiminnan prosessi- ja palvelukeskeiselle kehittämiselle. Kun palvelurajapinnoissa välitettävät tiedot on jäljitettävissä yritystietomallin käsitteisiin, palveluketjujen kehittäminen liiketoimintaprosessin muuttuessa tulee joustavammaksi ja nopeammaksi. Muita kokonaisarkkitehtuurin ja sen osana olevan informaatioarkkitehtuurin tyypillisiä tavoitteita löytyy Aija Palomäen artikkelista. Pystytkö tunnistamaan mille osa-alueille ja minkätyyppisten järjestelmien rakentajien tarpeisiin Nokian NEIM on tehty?

Kun yritystietomalli on saatu perustettua, jatkovaiheena voisi harkita tietojen merkitysten mallintamista käsitekarttoiksi, jotka kertovat käsitteiden suhteet eri asiayhteyksissä. Käsitekarttojen avulla voi tehokkaasti ja tarkasti välittää eri näkökulmiin ja osa-alueisiin liittyvää tietämystä ihmisten välillä. Jos käsitekartat mallinnetaan ontologioiksi semanttisen webin tekniikoilla, voidaan niitä käyttää systeemyön välineisiin integroituna apuna tietojen merkityksen näyttämisessä ja relevantin tiedon etsimisessä. Lue semanttisesta mallintamisesta ja ontologioista

lisää tämän lehden muista artikkeleista.



Anna Aaltonen työskentelee Capgeminissä tietotekniikkakonsulttina, toimenkuvanaan erilaiset määrittelytehtävät;
anna.aaltonen@capgemini.com



Jouni Lähteenmäki työskentelee Osuuspankkikeskuksessa arkkitehtuuri- ja kehitystehtävissä;
jouni.lahteenmaki@op.fi

Tietovaraston sovellusarkkitehtuuri muutoshallinnan näkökulmasta

Henry Runn, Tietollmarinen



Tässä artikkelissa kuvataan tietovaraston sovellusarkkitehtuuri, joka huomioi mahdollisimman kustannustehokkaasti liiketoiminnassa tapahtuvia muutoksia. Aluksi esitetään erilaisia tietovarastoratkaisuja, jonka jälkeen käydään läpi muutoshallinnan ja kustannustehokkuuden näkökulmasta tietovarastointiprosessi sekä tietomalli. Lopuksi nämä asiat nivotaan yhteen ylläpidettävyyden osalta. Tässä artikkelissa tietovaraston oletetaan olevan tietokantapohjainen. Tietovarastoinnilla tarkoitetaan sitä prosessia, jonka avulla tiedot poimitaan lähdejärjestelmistä (operatiivisia tai ulkoisia järjestelmiä), muokataan tietovarastoon tallennettavaksi sekä tarjotaan loppukäyttäjille käytettäväksi.

Tietovaraston merkitys liiketoiminnalle

Yritys voi ottaa arkkitehtuurissaan kantaa siihen, mikä on tietovaraston rooli informaation hallinnassa. Arkkitehtuurin mahdollisia näkökulmia tietovarastoon ovat:

- Tietovarastoa ei nähdä lainkaan tarpeellisena, jolloin yrityksellä on jo käytettävissään kaikki se informaatio, jonka tietovarasto voi tuottaa.
- Tietovarasto on tärkeä osa yrityksen liiketoimintaa, jolloin yritys kokee että eri järjestelmissä olevien tietojen yhdistä-

misellä on saavutettavissa liiketoiminnallista hyötyä.

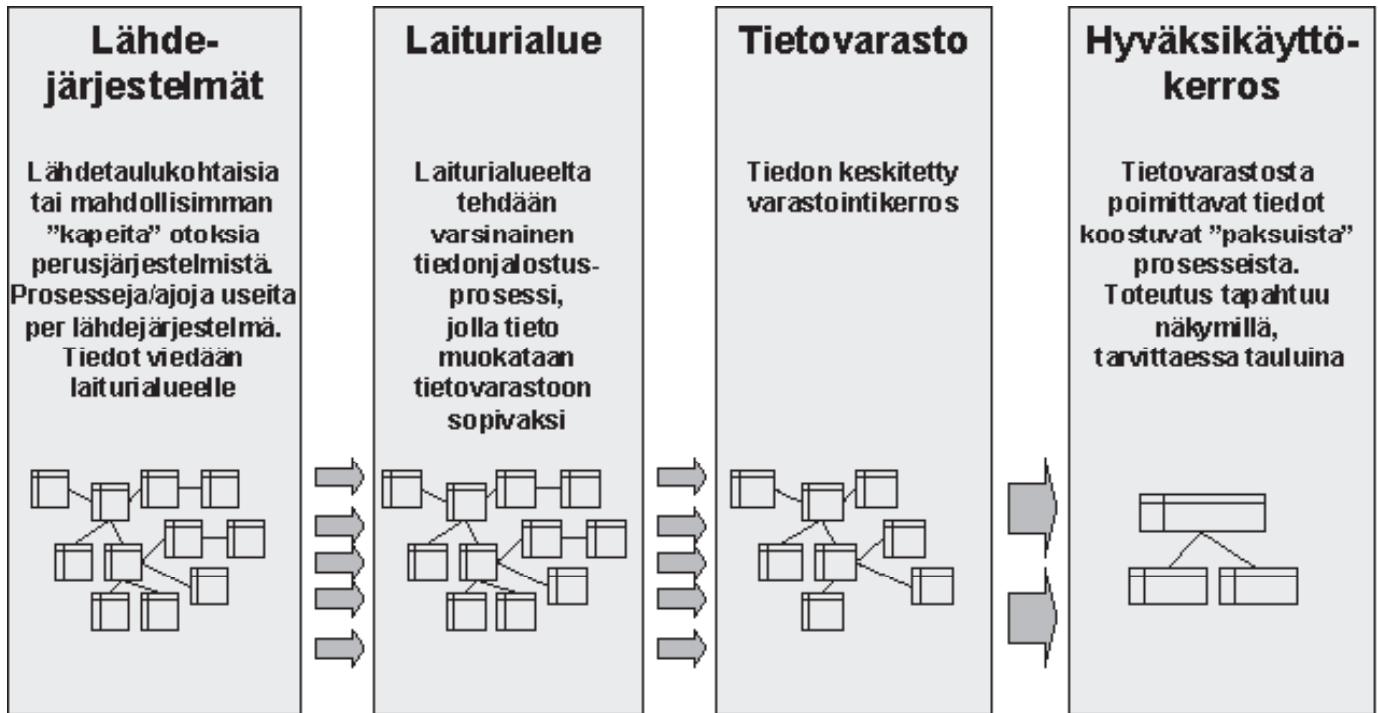
- Tietovarastolla on merkitystä vain osasto/linjakohtaisina ratkaisuna, jolloin yritys kokee että osastoilla käytettävien järjestelmien tietojen yhdistämisellä on saavutettavissa liiketoiminnallista hyötyä, mutta yli osastojen tapahtuva tietojen yhdistäminen ei tuota lisäarvoa.

Tietovarastointiprosessi muutoshallinnan näkökulmasta

Nykymaailmassa ainoa varmalta tuntuva asia on se, että asiat muuttuvat. Tietojärjestelmiä toteutetta-

essa on muutosten hallinta tärkeä osa onnistunutta toteutusta. Miten tietovarastosta saadaan muutoksia kestävä ja varsinkin sellainen, että muutokset voidaan tehdä nopeasti ja kustannustehokkaasti?

Muutostenhallinnan kannalta tietovarastointiprosessi on jaettava sopiviin osakokonaisuuksiin, joiden sisällä muutoksia voidaan toteuttaa häiritsemättä muita osakokonaisuuksia. Tietovarastointiprosessin jako osakokonaisuuksiin voidaan tehdä esimerkiksi seuraavalla tavalla:



Kuva 1. Tietovarastointiprosessi.

Prosessin ensimmäinen tehtävä on hakea/tuoda tiedot lähdejärjestelmistä tietovaraston laiturialueelle tallentaen ne samaa muotoon kuin tiedot ovat lähtöjärjestelmässä. Tämä laiturialue toimii paluupisteinä, jos seuraavat vaiheet tietovarastointiprosessissa epäonnistuvat, jolloin prosessia ei tarvitse aloittaa täysin alusta. Tämä on hyvinkin tärkeää tilanteissa, joissa lähdejärjestelmistä ei voida päiväkään lukea tietoja tietovarastoon. Epäonnistuneet prosessithan havaitaan yleensä aamulla, kun ihmiset tulevat töihin.

Seuraava osakokonaisuus on prosessi, joka siirtää tiedot laiturialueelta tietovarastoon. Tässä prosessissa tehdään huomattava määrä tiedonjalostusta ja muokkausta, joten muutosten hallinnan kannalta tämä prosessi kannattaa pitää mahdollisimman kapeana. Prosessin kapeudella tarkoitetaan sitä, että pyritään tekemään

mahdollisimman pieniä ja siten helposti hallittavia kokonaisuuksia. Tämän lisäksi prosessin avulla voidaan erottaa lähdejärjestelmässä tapahtuvat muutokset tietovarastosta eli päättää, mitkä tiedot tuodaan tietovarastoon ja mitkä jätetään tuomatta.

Viimeinen osakokonaisuus on prosessi, joka mahdollistaa tiedon hyväksikäytön tietovarastosta. Tietoturvan kannalta ei ole suositeltavaa, että käyttäjät pääsevät suoraan tietovarastoon. Heille voidaan tehdä joko virtuaalinen hyväksikäyttökerros tietovarastoon tai heidän tarpeitaan varten rakennetaan oma data mart tai näiden kahden yhdistelmä. Virtuaalinen hyväksikäyttökerros eli näkymä on erittäin joustava tapa tuottaa tiedot hyväksikäytettäväksi, koska se ei vaadi erillisiä tiedonsiirtoja.

Tärkein valintakriteeri hyväksikäyttökerroksen tuottamisessa on

nopeus. Jos näkymä toimii yhtä nopeasti tai lähes yhtä nopeasti kuin erillinen fyysinen taulu, on näkymä suositeltava ratkaisu. Jos näkymä toimii hitaasti, on olemassa kaksi vaihtoehtoa. Ensiksi näkymän toimintaa voidaan nopeuttaa indeksoimalla tai muodostamalla materialisoitu näkymä, jolloin myöskään ei tarvitse tehdä erillistä tiedonsiirtoa. Toinen vaihtoehto on taulun ja siihen liittyvän tiedonsiirtoprosessin luominen. Tämä ratkaisu on muutosten hallinnan kannalta hitain, sillä siihen sisältyy tiedon rakenteen esittämisen lisäksi tiedon tuottaminen.

Tietovaraston tietomalli muutoshallinnan näkökulmasta

Tietomallin fyysiseen toteutukseen pätevät samat säännöt kuin prosesseihin. Jos tiedot on tallennettu leveisiin taulurakenteisiin, niissä tapahtuvien muutosten hallinta on hankalampaa ja hitaampaa kuin kapeissa taulurakenteissa,

mikäli tauluun on koostettu tietoja useista eri lähteitä yhdistelemällä. Esimerkki tästä on dimensiomallin faktatiedot, jossa faktat on koostettu useasta eri järjestelmästä.

Toinen muutosten hallintaan vaikuttava tekijä on luonnolliset avaimet. Luonnollisissa avaimissa tapahtuvat muutokset pitää vierittää jokaiseen tauluun, jossa luonnollista avainta on käytetty. Luonnollisten avainten sijaan voidaan käyttää surrogaatteja. Tällöin luonnolliset avaimet tallennetaan vain yhteen tauluun, johon luonnollisissa avaimissa tapahtuvat muutokset kohdennetaan. Surrogaattien käyttäminen on muutosten hallinnan kannalta suositeltava ratkaisu.

Tietovarasto sisältää myös tietojen historiatiedon. Tiedon muuttumisen näkökulmaan ei usein kiinnitetä huomiota, kun tietovaraston fyysinen toteutus tehdään. Mikäli muuttumisen näkökulmaa käytetään fyysisen toteutuksen suunnittelussa, niin tiedot tallennetaan sellaisiin tietorakenteisiin, joiden muuttumissykli on sama.

Esimerkiksi, jos henkilön nimi vaihtuu todella harvoin verrattuna henkilön osoitteeseen, henkilön nimi ja osoite voidaan tallentaa eri tauluihin. Tämä tiedon muuttumisen näkökulman hyödyntäminen fyysisessä toteutuksessa vähentää tallennettavan tiedon määrää, koska siinä pyritään minimoimaan tiedon ylimääräistä tallennusta. Esimerkkinä voidaan ajatella edelleen asiakkaan nimeä ja osoitetta. Jos ne ovat samassa taulussa ja osoite vaihtuu viisi kertaa, asiak-

kaan sama nimi esiintyy taulussa viiteen kertaan. Mikäli ne ovat eri tauluissa, on nimi taulussa vain yksi rivi ja osoitetaulussa viisi riviä.

Muutoshallinnan merkitys tietovaraston kokonaiskustannuksiin

Kun tarkastellaan tietovaraston kokonaiskustannuksia pitkällä aikavälillä, voidaan ajatella, että infrastruktuuriin tehty investoinnit ovat vain osa kokonaiskustannuksista. Suurin kustannus lienee tietovaraston toteuttamiskustannus, mutta tietovaraston ylläpito aiheuttaa myös huomattavia kustannuksia, jotka ennen pitkää ylittävät toteuttamiskustannukset.

Ylläpitoon liittyvät kustannukset riippuvat valitusta arkkitehtuurista sekä tietovarastoinnin toteutustavasta. Jos arkkitehtuurissa on päädytty osastokohtaisiin tietovarastoihin, on niiden ylläpitokustannus suurempi kuin keskitetyn tietovaraston, mikäli yrityksellä on käytössään useita yhteiskäyttöisiä operatiivisia järjestelmiä, joihin on rakennettu osastokohtaiset liittymät yhden keskitettyyn tietovarastoon tehdyn liittymän sijasta. Tämä sillä edellytyksellä, että yhden liittymän ylläpitäminen maksaa vakiosumman.

Toteutustapa vaikuttaa ylläpidon kustannuksiin. Jos toteutus on tehty yhdellä prosessilla, pienikin muutos operatiivisessa / ulkoisessa järjestelmässä, tietovarastossa tai hyväksikäyttökerroksessa johtaa muutokseen koko prosessissa. Tämän pitkän ja paksun prosessin muuttaminen on kallista ja hidasta,

jos muutos koskee ainoastaan jotain osaa ylläkuvausta osakokonaisuuksista. Jos muutos koskettaa koko prosessia, niin muutoksen hallinta ja varsinkin sen testaus on helpompaa, koska muutos voidaan tehdä osakokonaisuus kerrallaan.

Artikkelin kirjoittaja Henry Runn työskentelee Tietollmarisessa toimenkuvanaan tietovarastointi, henry.runn@tietoenator.com

IIR:n seminaari Tehokas Business Tiedon hallinta 05

Osallistuin IIR:n seminaariin 'pressikortilla' syyskuussa kuullakseni, mitä suomalaisille yrityksille tarkoittaa Liiketoiminta tieto (BI) ja miten sitä hyödynnetään tehokkaasti. Esitykset olivat kaikki suunnattu ensisijaisesti markkinoinnin ammattilaisille.

Tämän hetken trendi näyttäisi näiden esitysten valossa olevan sellainen, että yritykset panostavat liiketoiminnan tarvitseman tiedon keruuseen ja jakelukanavan helpouteen. Tietoa kerätään prosessinohjaus-, markkina- ja kilpailijaseurannan näkökulmista. Tiedon jakelukanavana on käytössä sisäinen internet (tai erillinen sisäiseen käyttöön tarkoitettu BI -portaali), jolla tieto saadaan helposti jaetuksi. Postin tietopalvelupäällikkö Hannu Kyyrö kiteytti asian hyvin heidän määritelmällään BI:lle: "Systemaattinen toimintamalli, joka tuottaa ja viestii jäsenettyä ja analysoitua tietämystä liiketoimintaympäristöstä". On osattava löytää olennainen tieto ja tieto on helposti saatava oikeille henkilöille, kun he tietoa tarvitsevat.

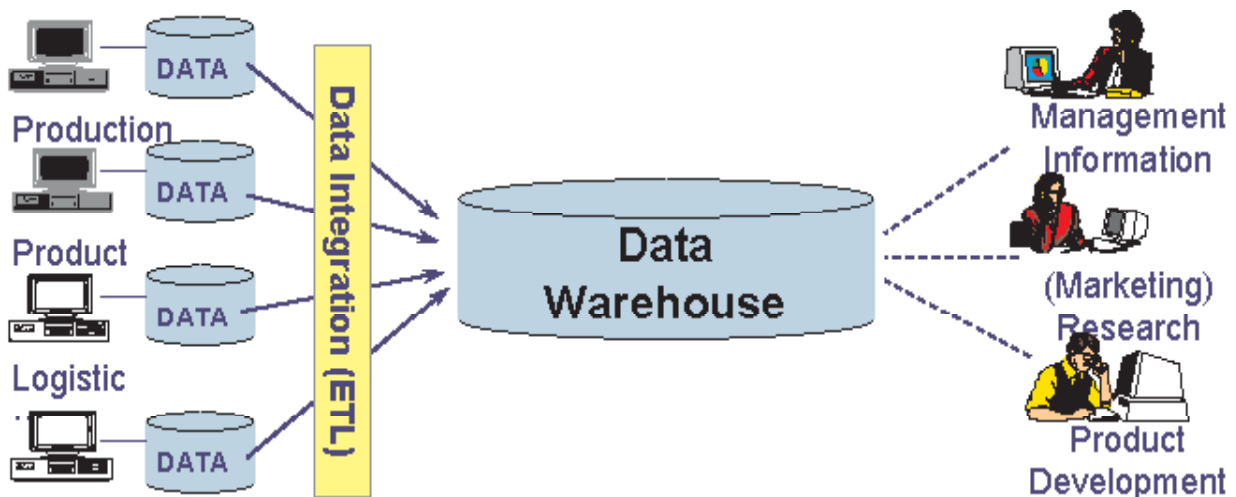
Minna Oksanen

Sirpaletiedosta kokonaisuuteen

Minna Oksanen, DAMA Finland vpj.



Artikkelissani käsittelen prosessia, jolla operatiivisten järjestelmien tietoalkioista saadaan aikaan integroitu kokonaisuus kohdentaen tiettyyn objektiin kuten asiakas. Prosessia voidaan verrata viininvalmistukseen, kuten eräs mainoskampanja aikoinaan kertoi. Operatiiviset järjestelmät tuottavat rypäleitä ja tietovarasto muodostaa viinin. Jotta viinistä saadaan hyvää, sen valmistukseen tarvitaan ohje ja sitä pitää säilyttää tynnyreissä. Ohjeena tietovarastoinnissa toimii mallinnus. Ensin on mallinnettava, mitä tietoa kohteesta tarvitaan, jotta tiedetään mitä tietoa tulee kerätä. Ennen kuin tiedot viedään tietovarastoon, ne pitää muokata yhteensopiviksi kuten viinin valmistuksessa, rypäleet murskataan ja lisätään muut tarvittavat aineet. Tietovarastoinnissa tästä vaiheesta käytetään nimitystä tiedon integraatio. Lähtökohtana on Bill Inmonin käyttämä kokonaisvaltaisen tietovarastointimallin määritelmä ETL- (Extract transfer and load) vaiheelle. Tietovarastolle kuten viinille on asetettava myös laatuvaatimukset. Jotta lopputulos on tarkoituksenmukaista on kiinnitettävä huomiota tiedon sisällölliseen ja tekniseen laatuun. Tietovaraston rakentaminen ei ole yksittäinen projekti vaan se on jatkuvaa kehittämistä. Tulevaisuudessa haasteeksi nousee se, kun tietojen analysointi tehdään entistä nopeammalla syklillä, miten tietomassat saadaan hallitusti käsiteltyä.



Kuva 1. Kokonaisvaltainen tietovarastointi arkkitehtuuri liiketoiminnan näkökulmasta

Kuten kuvasta 1. voidaan havaita, tiedonhallinnassa tarvitaan useita vaiheita. Operatiivisten järjestelmien tehtävänä on toimia tehokkaina tiedon syöttö- ja käsittelyalustoina. Tieto kohdennetaan tiettyyn toimintoon kuten tilausten ja laskujen käsittelyyn, kirjanpitoon ja varastonvalvontaan. Kun yritys tarvitsee tietoa, jota tuotetaan useassa järjestelmässä, tarvitaan yhtenäistä tietoa. Esimerkiksi yritys tarvitsee tietoa yksittäisen asiakkaan tai asiakasryhmän käyttäytymisestä. Kysymyksiä, joihin etsitään vastauksia, voivat olla esimerkiksi seuraavat. Montako tilausta asiakas on tehnyt? Montako toimitusta tammikuussa on suoritettu? Onko asiakkaalla avoimia laskuja, ja lopullisena kysymyksenä, joka liiketoiminnassa on tärkeintä, onko asiakas yritykselle kannattava? Asiakkaan näkökulmasta vastaava kysymys kuuluu: "Saanko hyötyä yrityksen tarjoamasta palvelusta?"

On kaksi vaihtoehtoista tapaa toteuttaa raportointi edellä mainitun kysymyksen selvittämiseksi. Voidaan tehdä jokaisesta järjestelmästä erikseen raportit, jotka yhdistetään esimerkiksi Exceltaulukkoan analyysiä varten. Voidaan myös tehdä raportti yhdestä lähteestä, johon tieto on kerätty asiakasnäkökulmasta. Jälkimmäisestä ratkaisusta käytetään nimitystä asiakastietovarasto ja edellisestä raporttikaaos. Useat eriliset raportointijärjestelmät voivat vääristää kuvan tilanteesta, laskentakaavat vaihtelevat ja lähtötietoja käsitellään eri tavoin. Kerättyä tiettyyn kohteeseen liittyvää tietoa pyritään saamaan "a single view of

the truth" eli yhtenäinen kuva. Kun hyödynnetään tiedot yhtenäistävää prosessiä tiedonkeruuseen ja muokkaukseen, saadaan tiedon loppukäyttäjille aina vastaava tieto.

Kokonaisvaltaisen tietovarastointi-idean isä Bill Inmon määrittelee asian seuraavasti: "A data warehouse is a subject-oriented, integrated, time-variant, non-volatile collection of data used to support the strategic decision-making process for the enterprise and is the central point of data integration for business intelligence." Vapaasti suomentaen: "Tietovarasto on kohteeseen sidottujen tietojen yhdistäminen eheäksi kokonaisuudeksi, jossa on huomioitu aikaperspektiivi ja jota hyödynnetään yrityksen strategisten päätösten teossa."

Koska tietovarastoa hyödynnetään strategiseen päätöksentekoon, parhaiten tietovarastohanke onnistuu, kun liiketoiminta riittävän korkealla tasolla on sitoutunut hankkeeseen jo heti sen alkuvaiheessa. Ennen kuin tietoja lähdetään 'haalimaan' operatiivisista ja muista mahdollisista lähteistä, tulee liiketoiminnan ja toteuttajien yhdessä päättää ne strategiset kysymykset, joihin halutaan vastaus.

Tiedot on mallinnettava liiketoiminnan tavoitteiden mukaan

Kun tiedetään tietovaraston tietotarve, on tiedot mallinnettava järkeväksi kokonaisuudeksi. Vasta tämän jälkeen voidaan aloittaa varsinainen tiedonkeruuprosessi. Tietovarastoterminä vaiheesta

käytetään nimitystä tiedon integrointi eli ETL-vaihe.

Tiedon integrointi on 80% työmäärästä

Ensimmäisellä tietovarastointiin liittyvällä kurssilla, jonka kävin 1997 SAS Instituutissa kuulin väitteen, että ETL-vaihe on 80% kokonaistyömäärästä. Raporttien ja analysointijärjestelmien toteutuksen osuudeksi jää vain 20%:a. Perinteisen operatiivisen raportoinnin tekijälle väite kuulosti melko uskomattomalta. Nyt noin 8 vuotta alan töitä tehynä olen täysin samaa mieltä. Joskus jopa tuntuu, että tiedon integrointi on 99%:a kokonaistyöajasta varsinkin, jos integrointivaiheeseen huomioidaan mukaan tietovaraston perustaminen, massojen ja tehokkuuden hallinta sekä jatkuva muutostarve. Itse raportit todellakin vain muodostetaan valmiista, muokattusta aineistosta. Mikäli liiketoiminnan tavoite on epäselvä, voi siitä aiheutua sisäistä politikointia ja jopa työvaiheiden viivästystä tai jarrutusta. Liiketoiminta on siis saatava sitoutetuksi prosessiin, jotta lopputuloksesta tulee kannattava.

Latausalue

Landingarea, stagingarea, eteinen, laitur, latausalue - rakkaalla lapsella sanotaan olevan useita eri nimiä ja niin on myös tietointegraation tärkeimmällä osalla. Koko ETL-prosessi tapahtuu latausalueella joka on looginen kokonaisuus. Fyysinen toteutus voi olla hajautettu. Latausalue voi olla oma erillinen tietokantansa tai osa tietovarastokokonaisuutta. Latausalue tarvitaan, koska ennen kuin tiedot

viedään lopulliseen varastoon, ne on yhdenmukaistettava. Tämä työ on tehokkainta tehdä ennen, koska yksittäisen latauksen tietomassa on helpompi hallita. Määrä on suppeampi ja keskimäärin aina lataus kohdistuu yhteen ajanhetkeen.

Ensin on E ja sitten T ja lopuksi L, Q ohjaa kaikkia.

“E stands for extract , T for transform and L for Load”, siis näistä kirjaimista muodostuu tiedon integrointivaiheen perusprosessi. Q eli quality on kaiken kattava ohjaava tekijä. On muistettava että tietovaraston sisällön arvo liiketoiminnalle on suorassa suhteessa sen tietosisällön laatuun.

Ensin tarvittavat tietoalkiot on kerättävä, haalittava ja poimittava sopivista lähteistä . Aina valmista tietoa ei ole edes saatavilla, vaan se on muodostettava joko uutena tietona lähdejärjestelmiin tai muokkausprosessin yhteydessä. Seuraavaksi eli tiedon muokkausvaiheessa eri lähteistä kootut tiedot yhdenmukaistetaan ja lopuksi ne ladataan tietovarastoon. Tietyillä latausvälineillä toimittaessa voidaan kääntää T ja L, koska on tehokkaampaa ensin ladata tiedot ennen kuin niitä lähdetään muokkaamaan.

E (etsi tarvittava data)

Tietojenkeruuvaiheen ensimmäinen haaste on, mitkä tiedot tarvitaan / vaaditaan mukaan. Määrittely- vaiheessa tietovaraston tietomallinkentät ‘mäpätään’ lähtöjärjestelmien tietoihin ja samalla päätetään käsittelysäännöt, joilla tiedot poimitaan. Poiminnan yhteydessä tehdään tarvittavat tietotar-

kistukset. Esimerkiksi: tarkistetaan asiakastunnuksen muoto, ovatko 0 ja tyhjä samanarvoisia, ovatko päivämäärät oikeita. Esimerkiksi: ladataanko tietovarastoon tietoa, joka on viisi vuotta vanhempaa? Voivatko kyseiset tapaukset olla nykytietojen korjauksia, vai ovatko ne vain virhetapauksia? Vastauksen voi kertoa vain tiedon omistaja eli liiketoimintayksikkö. On myös päätettävä mitä tehdään virheellisille tapauksille. Otetaanko ne mukaan lataukseen ja ilmoitetaan niistä virhelistalle, jätetäänkö ne käsittelystä vai odotetaanko operatiivisen yksikön korjaustoimenpiteitä.

Kun tiedetään, mitä tietoa tarvitaan, pitää tieto vielä saada siirretyksi. Keruuvaiheen tekninen haaste on lähdejärjestelmien laaja kirjo. Lähdejärjestelmiä on tyypillisesti useilla eri laitealustoilla; on kirjanpitoa, asiakasjärjestelm(i)ä ja kaikista näistä tulisi saada tiedot kerätyksi. Siirretäänkö tiedot peräkkäistiedostona, ftp-palveluna vai luetaanko lähdejärjestelmiä suoraan vai käytetäänkö XML-sanoman välitystä. Tiedonsiirrossa on mietittävä onko järjestelmien välinen yhteys jatkuvassa päivityksessä käytössä vai kuormitetaan siirtoa vain kuukausittain. Nämä päätökset vaikuttavat teknisen ratkaisun valintaan. Laatu- ja turvallisuuskulmasta keruuvaiheen tehtävä on valita ja toimittaa tarkoituksen mukainen tieto jatkokäsittelyyn.

T (tuota informaatio)

Tietojen muokkaus (transform) käsittää eri tietolähteistä saadun tiedon yhdistämisen, muokka-

uksen ja puhdistamisen, joita kutsutaan yhteisnimellä clensing. Tunnukset, koodit ja mittayksiköt on yhdenmukaistettava, jotta päästään tavoitteeseen eli yhtenäiseen kuvaan kohteesta. Toisessa järjestelmässä voidaan käyttää 1 ja 0 kertomaan kyllä / ei tietoa, kun toisessa vastaavat arvot ovat Y ja N. Tiedon eheys on varmistettava. Leski- ja orporivit tulee estää, jos ne ovat mahdottomia. Voiko olla tilauksia, joilla ei ole asiakastietoa? Voi olla, että tilausjärjestelmän tiedot päivittyvät nopeammin kuin uusi asiakas saadaan perustettua. On myös tehtävä käsittelysäännöt, kun sama tieto voi tulla useasta lähteestä. Joissain tilanteissa käsittelyn helpottamiseksi on mietittävä pitääkö tietoa monistaa. Toteutuksen kannalta muokkausvaihe on työläin. Jokainen muokkaussääntö on koodattava ja ne on myös pystyttävä testaamaan.

L (lataa tietovarastoon)

Lataustapa on asia, joka on päätettävä määrittelyn yhteydessä. Onko tarve ladata aina kaikki tiedot vai riittääkö, että tuodaan uudet ja muutetut tiedot? Mitä tehdään, jos tieto poistuu, historioidaanko vanha vai hävitetäänkö kokonaan pois? Jos halutaan tehdä aikasarja raportointia kahden vuoden ajalta, tarvitaanko tällöin kolme, neljä vuotta vanhempaa tietoa? Kun asiakas muuttaa ja osoite vaihtuu, tarvitaanko aikaisempaa osoitetta? Kun tuoteluokittelu muuttuu joka kuukausi, pitääkö ladata kaikki aikaisemmat luokitukset? Nämä ovat päätöksiä, joihin vaikuttaa taas liiketoiminnan vaatimukset ja tiedon käyttötarve. Itse lataus kuulostaa yksinkertaiselta toimin-

nolta. Painetaan nappia ja kone siirtää tiedot tietokantaan. Entäpä, kun kanta on indeksoitu, voiko indeksoituun kantaan ladata suoraan tietoa? Miten varmistetaan, että kaikki tiedot on ladattu? Lisää kysymyksiä, joihin vastauksen pystyvät antamaan tietokantaa tuntevat henkilöt. Eri kantaratkaisissa toiminta vaihtelee.

Q (kvantitatiivinen ja kvalitatiivinen laatu)

Laatuvaatimukset voidaan luokitella sekä teknisten että sisällöllisten laatuksien mukaan. Teknisiä ovat määrälliset ja muodolliset tekijät. Kuten se, että kun tietoa siirretään laitealustalta toiselle, on varmistettava, että kaikki tieto siirtyy. Muodollisella laadulla tarkoitetaan esimerkiksi päivämäärä -kentän muotoa, voidaanko numeerista ja merkkimuotoista tietoa verrata, onko tietoalkioiden pituus riittävä, ettei esimerkiksi summauksen yhteydessä tapahdu ylivuotoa.

Sisällölliset laatuksien kriteerit ovat vaatimuksia tietojen oikeellisuudelle. Laadun oikeellisuuden taso on sovittava tietovaraston ja siihen tietoa tuottavien lähdejärjestelmien omistajien välillä. Olisi hyvä olla osapuolet sitova sopimus, jossa sovitaan toimintatapa lähde-

järjestelmien ja tietovaraston välillä tapahtuvalle tiedonsiirrolle. Sopimuksessa tulee käydä ilmi, mitkä tiedot kukin lähdejärjestelmä on sitoutunut toimittamaan sovittuna ajanhetkenä sekä siinä on sovittava toimenpiteistä poikkeustilanteissa.

Jotta tietovaraston tavoite ”Single view of the truth” ilman ”crab in crab out” ominaisuutta, tiedoille pitää määrittää laatuvaatimukset. voidaan löytää useita eri vaatimuksia sekä määrällisiä ja laadullisia että teknisiä ja sisällöllisiä laatuvaatimuksia. Tavoitetaso tarvitaan, koska täydellistä laatua ei voi saada ilman kustannuksia. Tietovarastossa on myös olemassa tietoa, jonka laadusta ei voida olla varmoja. Kaiken kaikkiaan laatua on raportoitava ja mitattava.

“Metadata is the glue and clue”

Integraation prosessin kuten koko tietovarastoinnin hallinnointi vaatii tietoa prosessista ja tiedoista eli metadataa. Metadataa on teknistä, liiketoiminnallista ja hallinnollista. Oheinen kuva kuvaa, miten liiketoiminnan vaatimusten pohjalta syntyy käsittelysäännöt tiedon muokkaukseen. Näistä käsittelysäännöistä muodostuvat liiketoiminnallinen metadata. Tekaninen ja hallinnollinen meta-

data syntyy toiminnallisuusvaatimuksista.

Teknisellä metadatatalla integraatioprosessin yhteydessä tarkoitetaan tietojen fyysisen rakenteen kuvausta, joita ovat taulut, lähtötiedostot ja tietoalkioiden muodot. Lähtöjärjestelmissä tiedot on kuvattu ja tänä päivänä on mahdollista hyödyntää suoraan OMA- (Open Metadata Architecture) standardin avulla lähtöjärjestelmän kuvauksia tietovarastossa. Standardi pohjautuu XML-muotoon.

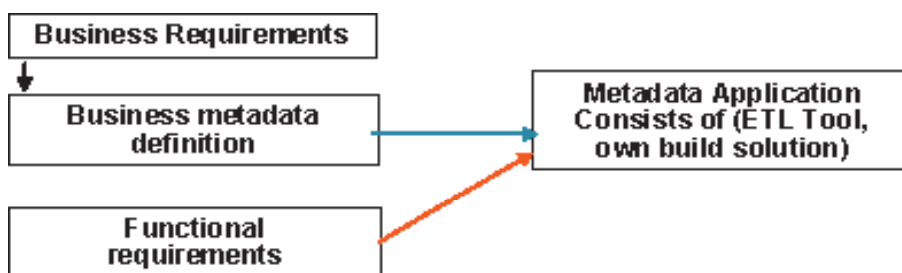
Markkinoilla on muutama johtava Data Integraatioon kehitetty työväline

Tiedon integraatioprosessia on helpotettu eri tuotetoimittajien taholta kehittämällä prosessiohjautuvia työvälineitä. Yhtenäisellä välineellä, joka on metadata ohjautuva, kuvataan tiedot ja muokausprosessi. Lopuksi kuvauksen mukainen prosessi generoidaan lähdekoodiksi. Lähdekoodi on usein SQL – pohjainen, jolloin metadata on siirrettävissä ja useimmiten ylläpidettävät tietovarastot ovat relaatiokantoja.

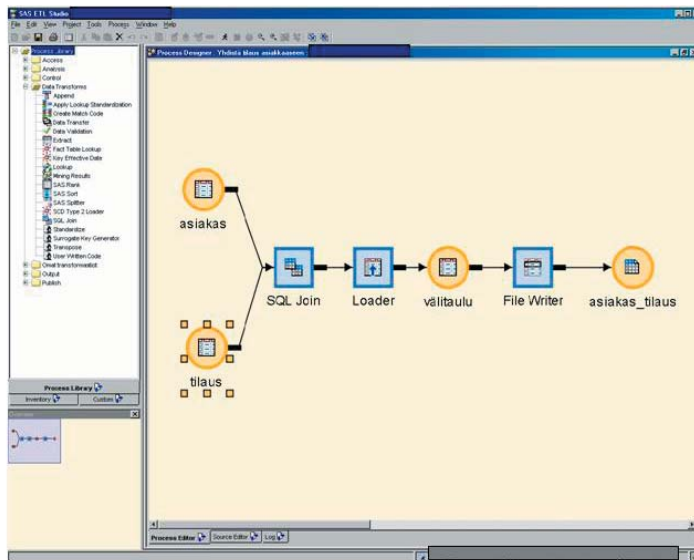
Haasteita

Eräajoista sanoman välitykseen

Uusin haaste on se, että tietoja halutaan raportoida nopeutetulla syklillä esimerkiksi web-sivuston käytettävyyden seuranta tunti ja minuuttitasolla. Päivätason raportointi on helppo järkeistää, lataukselle on riittävästi aikaa, sillä integrointivaihe voi toiminnasta ja poimittavien lähteiden määrästä riippuen kestää jopa useita tunteja.



Kuva 2. Metadatan muodostusprosessi



Kuva 3. ETL-prosessikuvaus SAS:n ETL-Studiossa.

Hyvä esimerkki on suomalaisen päivittäiskaupan päiväraportointi. Kuittirivitasen tieto (noin 1 milj. riviä) ladataan joka päivä tietokantaan aamulla muiden eräajojen päätyttyä, jotta liiketoiminta voi tehdä raportteja heti aamupäivällä. Edelleenkin kuitenkin puhutaan eräajomuotoisesta latauksesta. Kun siirrytään tosiaikaiseen tietovaraston lataukseen, puhutaan yksittäisistä sanomamuotoisista tietojen latauksista. Tässä xml-muoto on avannut uusia mahdollisuuksia.

Yritykset toimivat globaalisti

Toinen haaste on kansainvälisyys. Yritykset eivät toimi vain yhdessä toimipisteessä, vaan toimintaa on eri puolilla maailmaa. Ongelmana on, milloin tiedot voidaan ladata häiritsemättä tietojen hyväksikäyttöä. Hyvänä esimerkkinä on myynnin seuranta päivätasolla. Kun Aasiassa on iltapäivä, Suomessa on aamu ja Atlantin toisella puolella vielä nukutaan. Ruotsalainen kollegani, joka on ollut mukana globaalissa hankkeessa, kertoi, että heillä oli vastaavassa tilanteessa 2 tuntia vuorokaudessa aikaa saada tiedot ladatuksi, jotta ne olivat

kaikkien käytössä oikeaan aikaan. Miten tässä tilanteessa varmistetaan tiedon laatu? Latausta ei voida sen kaatuessa suoraan uusia, vaan korjaukset on tehtävä myöhemmin. Tai sitten valitaan menetelmä, jossa tiedot ladataan kantaan sellaisenaan ja tarkistukset tehdään suoraan kantaan.

Lainsäädännön asettamat vaatimukset

Kolmas haaste on lainsäädäntö. Henkilötietosuojalaki antaa ohjeet, mitä tietoa mm. asiakkaasta voidaan tallentaa. Edellä mainitussa päivittäiskaupan tietovarastossa on ETL-vaiheessa tehty kolmannen osapuolen toimesta tarvittavat salaukset henkilö- ja kanta-asiakastunnuksiin. Näin varmistetaan, että kenenkään henkilön yksityisyys ei ole vaarassa. Vastaavasti telealan lainsäädäntö ohjaa, mitä tietoa voidaan ja pitää tallettaa puhelusta. Toinen näkökulma tietojen tallentamisvaateisiin on se, että yhdistelemällä tietoa voidaan selvittää väärinkäytöksiä. Erityisesti luottokorttiyhtiöt hyödyntävät tätä ominaisuutta. Jos henkilöllä on samanai-

kaisesti tapahtumia Helsingissä ja Barcelonassa, niin voidaan olettaa, että kortista on otettu kopio.

Muistitaulu

Yhteenvetona haluan kiteyttää viisi omasta mielestäni tärkeää asiaa, jotka varmistavat tiedon integroinnissa parhaan mahdollisen lopputuloksen:

- Liiketoiminnan sitoutuminen;
- Selkeä tavoite, miksi tietoa kerätään;
- Tiedot on mallinnettava ennen kuin tietoa lähdetään keräämään;
- Tiedon laadun varmistamiseksi on tehtävä sopimus laatutasosta kaikkien osapuolten välillä ja laatua on mitattava, ja
- Tietovarastoa ei tehdä yhdessä projektissa vaan sen rakentaminen on jatkuva prosessi.

Lähteet

Lähteenä käytetty Bill Inmonin kokonaisvaltaiseen tietovarastojärjestelmään pohjautuvia AEL:n tietovarastointi kurssien koulutusmateriaaleja ja eri projekteissa koettuja omia sekä entisen kollegani Sebastian Green'in kokemuksia. Sebastian on nykyisin Ruotsin NCR:llä Teradata asiantuntijana. Kiitos myös kollegoilleni Paula Järvenpäälle ja SAS Institutin DW-asiantuntija Tapio Kalmille hyvistä kommentteista.

Minna Oksanen on toiminut DW-aiheisissa hankkeissa kahdeksan vuotta pääsääntöisesti riskienhallinnassa, viranomaisraportoinnissa ja analyttisessä CRM:ssä. Hän on toiminut kouluttajana ja osallistunut kansainvälisiin projekteihin. Minna vaihtaa 1.1.2006 Deloitteen.



Tiedon elinkaari hallintaan

Päivi Happonen, Kansallisarkisto

Julkishallinnon ja yksityissektorin organisaatioiden toimintaympäristö sähköistyy koko ajan: yhä uusia tietojärjestelmiä otetaan käyttöön ja järjestelmäintegraatiot yleistyvät. Sähköisessä toimintaympäristössä asiakirjallisten tietojen elinkaaren hallinta ja ohjaaminen sekä tietojen arkistointi edellyttävät suunnitelmallisuutta ja asiakirjahallinnollisten ominaisuuksien toteuttamista tietojärjestelmissä. Näiden ominaisuuksien avulla pystytään turvaamaan asiakirjallisen tiedon eheyteen, alkuperäisyyteen, luotettavuuteen ja käytettävyyteen liittyvät vaatimukset. Asiakirjahallinnolliset ominaisuudet kytketään osaksi tietojärjestelmien toiminnallisuutta, jolloin tietojärjestelmät tukevat niihin sisältyvien tietojen luotettavaa ja aukotonta käsittelyä koko niiden elinkaaren ajan, varmistavat asiakirjallisten tietojen käytettävyyden ja säilymisen sekä mahdollistavat säilytysarvon määrittelyn ja tarpeettoman aineiston hävittämisen.

Asiakirjallisen tiedon elinkaarta hallitaan ja ohjataan pitkälti meta-tietojen avulla. Jotta edellä luetellut asiakirjallisen tiedon laatuvaatimukset sähköisessä ympäristössä voidaan toteuttaa ja taata, järjestelmän metatietomallin on tarkoituksenmukaista perustua joko arkistolaitoksen SÄHKE-metatietomäärittelyyn tai asiakirjojen kuvailun ja hallinnan metatiedoista annettuun julkisen hallinnon suositukseen (JHS 143). Arkistolaitoksen metatietomäärittely on ensisijaisesti suunnattu asiankäsittelyjärjestelmiin, mutta se on sovellettavissa myös dokumenttienhallintajärjestelmiin. JHS 143

kohdentuu nimenomaan muihin tietojärjestelmiin kuin asiankäsittelyjärjestelmiin.

Metatietomäärittelyt sisältävät runsaasti metatietoelementtejä ja niiden tarkenteita. Luotettavan tiedonhallinnan näkökulmasta on ensiarvoisen tärkeää, että suurin osa vaadittavista metatietoelementtien arvoista tallentuu järjestelmään automaattisesti; käyttäjän tallentamat arvot on rajattava minimiin ja niiden tallentamista on tarkoituksenmukaista ohjata erilaisin valintalistoin. Esimerkiksi dokumenttienhallintajärjestelmässä metatietoelementin ”asia-

kirjatyyppi” arvoa käyttäjä ei voisi vapaasti valita, vaan valittavissa olisivat ainoastaan ne asiakirjatyypit, jotka liittyvät käsiteltävään prosessiin.

Julkishallinnon organisaatioissa erinomaisena tietojärjestelmien metatietoarvojen lähteenä toimii järjestelmien taustalle integroitava arkistonmuodostussuunnitelma. Lisätietoja arkistonmuodostussuunnitelmasta metatietojärjestelmänä saa arkistolaitoksen SÄHKE-hankkeen kotisivuilta (www.narc.fi/sahke) sekä vuoden 2006 alussa julkaistavasta arkistolaitoksen asiankäsittelyjärjestelmiä

ja sähköistä arkistointia koskevasta määräyksestä.

Tieto kytkettävä prosessiin

Tietojärjestelmissä käsiteltävät tiedot ja asiakirjatiedostot on kytkettävä aina laajempaan kokonaisuuteen. Dokumenttienhallinta järjestelmien kansiorakenteen tulisi perustua organisaation toimintaprosesseihin, ei organisaatioyksiköihin eikä asiakirjatyyppeihin. Jos kansiorakenne tehdään asiakirjatyypiperusteisesti, tallennetaan kaikki sopimukset yhteen kansioon, kaikki tarjoukset toiseen kansioon ja kaikki muistiot kolmanteen kansioon. Tällöin yksittäisen muistion yhteys siihen asiaan, mihin se liittyy sekä muihin samaan asiaan liittyviin asiakirjoihin, katoaa. Kuvattu tilanne ei ole kestävä sähköisen arkistoinnin näkökulmasta. Asiakirjalliselle tiedolle määritellyt laatuvaatimukset täyttyvät vain, jos asiakirjan asiayhteys, prosessikytkös, säilyy.

Vastaava tilanne syntyy, kun sähköpostiviestit säilytetään joko sähköpostijärjestelmässä tai sähköpostien arkistointiin kehitetyissä järjestelmissä. Viestit ovat irti siitä käsittelyprosessista, mihin ne liittyvät. Sähköpostijärjestelmä tai sähköpostiarkisto ei ole dokumenttien- eikä asianhallintajärjestelmä.

Asiakirjallisten tietojen käsittelyyn liittyvät työnnkulut on sidottava organisaation toimintaprosessien kehittämiseen. Järjestelmään on tallennuttava kattavat merkinnät tietojen käsittelyvaiheista. Aukoton käsittelyketju ja automaattiset tilasiirtymät takaavat omalta osaltaan järjestelmän tietosisällön luotet-

tavuuden, eheyden ja alkuperäisyyden. Asiakirjallisten tietojen tilasiirtymät määrittävät prosessikuvausten yhteydessä: miten asiakirjallisten tietojen tilat muuttuvat prosessin edetessä, missä vaiheessa asiakirjaan ei enää voi tehdä muutoksia? Tilasiirtymät automatisoidaan, jolloin prosessiin sisältyvän ennalta määritellyn toiminnallisuuden toteuttaminen järjestelmässä muuttaa asiakirjatiedoston tilaa.

Loki- ja muutoshistoriatietojen tallentuminen

Tietojärjestelmän on automaattisesti tallennettava järjestelmän tietosisältöön kohdistuneet tapahtumat ja toimenpiteet lokitietoihin. Esimerkiksidokumenttienhallintajärjestelmässä metatietoihin tehtävien merkintöjen ja muutosten sekä asiakirjallisten tietojen käsittelyyn liittyvien toimenpidetietojen on tallennuttava lokitietoihin. Asiakirjatiedoston sisältöön kohdistuvista muutoksista ja muutosten tekijöistä järjestelmän on kerättävä säilytyshistoriatieto. Loki- ja muutoshistoriatietojen tallentuminen ja niiden seuranta ja valvonta vastaavat omalta osaltaan asiakirjallisille tiedoille määriteltyihin laatuvaatimuksiin. Nämä vaatimukset vaikuttavat myös siihen, kuinka kauan lokitietoja järjestelmissä säilytetään.

Tietojärjestelmän hävitystoiminnallisuus

Nykyisten tietojärjestelmien määrittelyvaiheessa asiakirjallisen tiedon elinkaaren loppupää on valitettavan usein unohtunut. Järjestelmät palvelevat luonnollisesti organisaation operatiivista

toimintaa, mutta toiminnan rationalisoinnin ja kustannustehokkuuden näkökulmasta pitäisi olla luontevaa, että sähköisiin tietojärjestelmiin sisältyvät tiedot voidaan säilyttää luotettavasti sähköisessä muodossa ja että tarpeettomaksi käyneet tiedot voidaan järjestelmästä hävittää.

Käytännössä tämä tarkoittaa sitä, ettei tietojärjestelmissä tuotettujen ja tietojärjestelmään tallennettujen tietojen säilyttämistä eikä hävittämistä ole juurikaan mietitty. Hävittäminen toteutetaan eräajoina siten, että kaikki tiettyä vuotta vanhemmat tiedot poistetaan järjestelmästä. Tämä ei vastaa asiakirjahallinnollisia vaatimuksia, sillä organisaation asiakirjallisten tietojen säilytysajat on määriteltävä prosessiperusteisesti: kaikkia muistioita ei säilytetä yhtä kauan, vaan säilytysaika päätetään niiden tietosisällön mukaan.

Metatiedot ohjaavat myös asiakirjallisten tietojen hävittämistä. Tämä edellyttää sitä, että järjestelmään tallennetuille tiedoille on etukäteen määritelty konkreettiset säilytysajat ja säilytysajan päättymisen laskentaperusteet. Laskentaperusteet sidotaan järjestelmän toiminnallisuuteen. Määritetyn toiminnallisuuden toteuttaminen tallentaa automaattisesti säilytysajan päättymisajan kohdan asianomaiseen metatietokenttään. Järjestelmän hävitettäviksi esittämät tiedot on jonkun henkilön aina hyväksyttävä.

Päivi Happonen työskentelee Kansallisarkiston asiakirjahallintoyksikössä ylitarkastajana, paivi.happonen@narc.fi.

Arkistolaitoksen sähköinen arkistointi- ja tietopalvelujärjestelmä

Markus Merenmies, Kansallisarkisto

Arkistolaitoksessa on kehitetty Sähke-hankkeessa ratkaisuja sähköisten asiakirjojen muodostamiseksi siten, että ne voidaan myös pysyvästi arkistoida ja siirtää arkistolaitokseen sähköisessä muodossa. Luonnollisena jatkona tälle on nyt käynnistetty kehittämishanke valtakunnallisen sähköisen arkistointi ja tietopalvelujärjestelmän rakentamiseksi julkishallinnon käyttöön.

Sähköisen aineiston muodostaminen

Arkistolaitoksessa on ollut käynnissä vuodesta 2001 Sähke-hanke, jossa tavoitteena on ollut, että sähköisesti muodostuva asiakirja-aineisto kyettäisiin aikanaan arkistomaan yksinomaan sähköisessä muodossa arkistolaitokseen. Hankkeessa on painotettu aineiston muodostamiseen liittyviä vaatimuksia, kuten metatietojen muodostamista ja asiankäsittelyprosessin kykyä taata aineiston autenttisuus ja todistusvoima myös pitkälle tulevaisuuteen. Yhtenä merkittävänä kohtana vaatimuksissa on kuvaus siitä, kuinka sähköinen aineisto, eli asiakirjat sekä näihin liittyvät metatiedot niin asioilta kuin toimenpiteiltä, siirretään sähköiseen päätearkis-

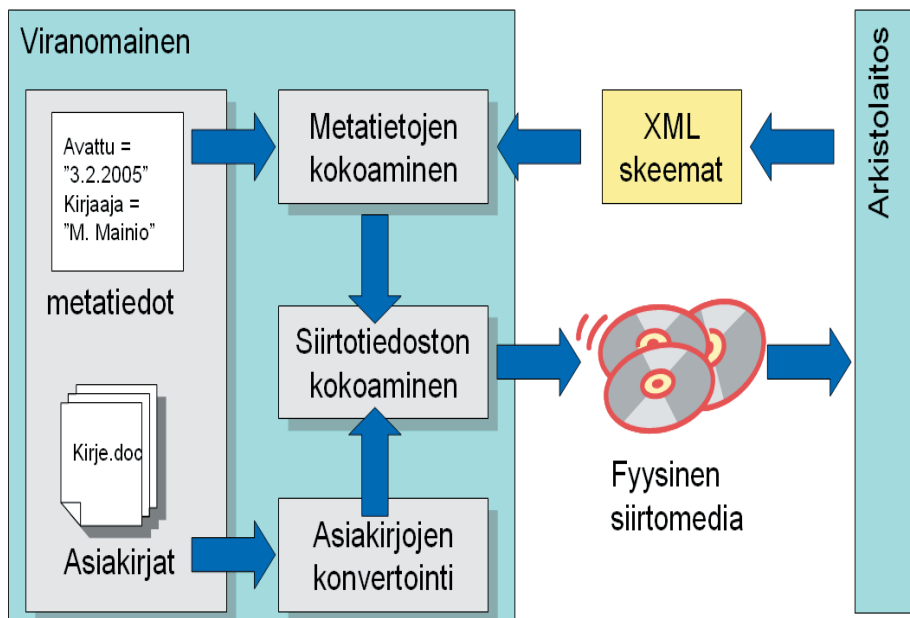
toon. Arkistolaitoksessa päädyttiin aikanaan siihen ratkaisuun, ettei arkistoida asiankäsittelyjärjestelmiä, vaan arkistointi kohdistuu yksinomaan niissä olevaan asiakirjalliseen tietoon. Tämän ansiosta on voitu kehittää Sähke-hankkeessa eri asiankäsittelyjärjestelmiä koskevat yhtenäiset vaatimukset aineiston muodostamiselle ja siirtämiselle sähköiseen arkistoon.

Siirtotiedosto, jolla asiat sekä asiakirjat siirretään, muodostetaan määräajoin asiankäsittelyjärjestelmästä. Paperiaikakaudella arkistolaitokseen siirrettiin asiakirjat vasta 40 vuoden jälkeen, mutta sähköisellä aikakaudella siirtosykliä on lyhennettävä, jotta aineiston säilyminen voidaan varmistaa. Lähtökohtana siirtymiselle sähköi-

seen arkistointiin on se, että asiankäsittelyjärjestelmä sisältää Sähkemäärittysten mukaisen metatietomallin ja että näihin aineistoa kuvaileviin tietokenttiin on myös vaadittavat tiedot tallennettu.

Sähköisten asiakirjojen arkistointi

Asiakirjojen säilyttäminen tekstin käsittelyohjelman omassa tiedostomuodossa on osoittautunut mahdottomaksi. Markkinoilla on tälläkin hetkellä, eräiden tuotteiden laajasta levinneisyydestä huolimatta, tarjolla lukuisia erilaisia tekstoreita ja näistä erilaisia versioita. Nämä teksturit ja versiot eivät tälläkään hetkellä ole yhteensopivia. Miltä tilanne näyttäisi 50 vuoden kuluttua kun näistä mainituista tekstoreista joudutaan



Kuva 1. Siirtokokonaisuuden muodostaminen

huomioimaan myös eri kehityssukupolvet? Arkistolaitos on tässä ottanut selkeän kannan siihen, että asiakirja-tiedostoja ei voida museoida alkuperäisessä muodossa, vaan ne on konvertoitava yhte-näiseen tiedostomuotoon, jonka avulla asiakirjan semant-tinen eheys voidaan taata. Semanttisella eheydellä tarkoitetaan sitä, että asiakirjan teknisen muodon sallitaan muuttuvan, mutta tietosi-sällön täytyy pysyä samana. Sähke-määritysten mukaan asiakirjat on konvertoitava kuviksi, eli TIFF-tiedostoiksi. Käytännössä tiedos-tomuoto on sama, kuin mitä käyte-tään asiakirjojen skannaamisen yhteydessä, mikä omalta osaltaan tuo yhtenäisyyttä tietojen käsit-telyyn. Konvertointi suositellaan tehtäväksi mahdollisimman aikai- sessa vaiheessa. Käytännössä se on parasta tehdä silloin, kun asiakirja asiankäsittelyjärjestelmässä lopul- lisesti lukittuu, eikä siihen enää sallita tehdä mitään muutoksia. Konversio on tehtävä viimeistään siinä vaiheessa, kun siirtotiedosto muodostetaan.

Yhtenä mielenkiintoisena keskus- telun aiheena on ollut PDF- tiedostomuodon käyttö arkistoin- nissa. Tässä on kuitenkin havaittu arkistoinnin näkökulmasta lukuisia ongelmia, kuten esimerkiksi PDF- tiedostoon sisältyvät makrot ja ettei tiedosto ei aina sisällä kaikkia tekstissä käytettyjä fontteja, jolloin asiakirjan lukeminen voi pahim- massa tapauksessa tulevaisuudessa estyä. Syyskuussa julkaistun PDF/ A-standardin on luvattu korjaavan nämä arkistolliset puutteet. Arkistolaitos selvittääkin parhail- laan niitä vaatimuksia ja teknisiä reunaehtoja, joita kyseisen tiedos- tomuodon mahdollinen hyväksy- minen edellyttää.

Sähköisen aineiston vastaanotto- ja palvelujärjestelmä

Aineiston siirto arkistolaitokseen

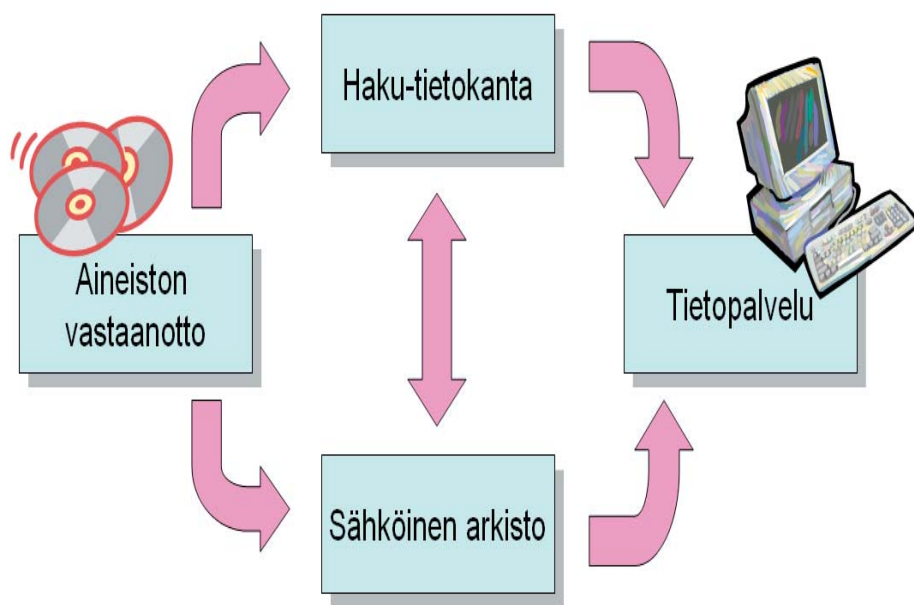
Aineisto siirretään viranomaiselta arkistolaitokseen käyttäen siirto- tiedostoa, joka sisältää niin meta- tiedot kuin varsinaiset asiakirjat.

Tämä siirtokokonaisuus voidaan tallentaa arkistolaitoksen hyväk- symälle tietovälineelle, kuten esimerkiksi CD-R. Sähke-määri- tykset kattavat aineiston muodos- tamiseen sekä siirtämiseen liittyvän ohjeistuksen, se mitä tapahtuu aineiston vastaanotta-misen yhte- ydessä ja sen jälkeen arkistolai- toksesta, on parhaillaan kehitteillä VAPA-hankkeessa (Vastaanotto ja Palvelu).

Siirtotiedoston muodostaminen on eräänlainen keräilytoiminto, joka kerää asiankäsittelyjärjestelmästä ne asiat, jotka ovat määrätyn ikäisiä. Asioihin ja asiakirjoihin liittyvät metatiedot sekä asioihin liittyvät asiakirjat tallennetaan arkistolai- toksen vaatimusten mukaisessa muodossa, eli metatiedot XML- tiedostoihin ja asiakirjat konver- toituina kuviksi.

Sähköisen arkiston rakenne ja tekniikat

Arkistolaitoksen sähköisen asia- kirja-aineiston vastaanotto- ja tietopalvelujärjestelmä koostuu seuraavista loogisista kokonai- suuksista: vastaanottojärjes- telmä, hakujärjestelmä, sähköinen arkisto sekä palvelujärjestelmä. Vastaanottojärjestelmän tehtävänä on tulkita siirrettävät aineistoko- konaisuudet ja avata siirtotiedot siten, että niihin sisältyvät meta- tiedot sekä asiakirjat voidaan arkis- toida oikeaan tietorakenteeseen. Metatiedot tallennetaan kahteen paikkaan, ensinnäkin hakujärjes- telmään, jossa on tieto kaikista arkistolaitokseen arkistoiduista aineistoista, siis myös paperimuo- toisista, sekä varsinaiseen sähköi- seen arkistoon. Hakujärjestelmään



Kuva 2. Sähköisen arkiston looginen rakenne

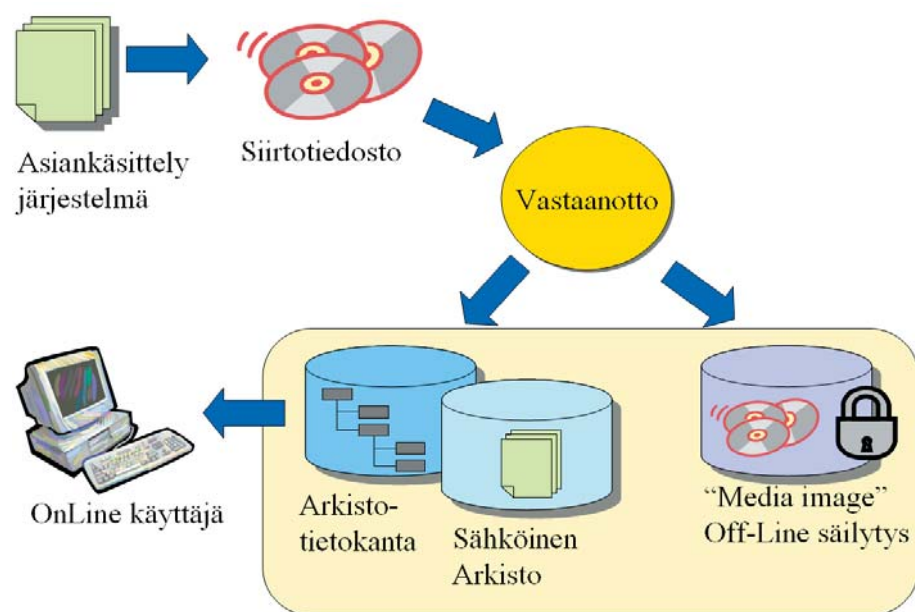
tallennetuilla metatiedoilla on tärkeä merkitys aineiston haun kannalta, jotta aineisto voidaan myöhemmin löytää osana arkistonmuodostajan arkistoa ja toimii näin linkkinä sähköiseen arkistoon, joka vastaavasti sisältää sekä metatiedot että asiakirjat. Koska julkishallinto on rakennettu julkisuusperiaatteen varaan, on tietopalvelun rakentaminen välttämätöntä, jotta tiedot saadaan niitä tarvitsevien käyttöön.

Sähköinen säilyttäminen erilaisilla siirtoon käytetyillä medioilla, kuten DVD tai CD-r, on aineiston ikääntymisestä aiheutuvan korruptoitumisen myötä erittäin epävarma ratkaisu, joka edellyttäisi aineiston mekaanista tuoreuttamista noin viiden vuoden välein. Tämän vuoksi arkistolaitos onkin valinnut tekniseksi ratkaisumalliksi yksittäisten mediataltoiden sijaan dataholvi-tekniikan, jossa tallennusratkaisu on levyjä ja nauhoja sisältävä suurikapasiteetinen järjestelmä. Kun arkistoitava

tieto tallennetaan tämänkaltaiseen aktiiviseen arkistointijärjestelmään, voidaan tietoa automaattisesti turvata erilaisia uhkia vastaan käyttäen tiedon tuoreuttamista, monistamista sekä maantieteellistä hajauttamista. Asiakirjallisen aineiston todistusvoiman säilyttämisen vuoksi on kiinnitettävä huomiota erityisesti eheyden ja alkuperäisyyden takaamiseen. Jotta aineiston muuttumattomuus

arkistossa voidaan varmistaa, on päädytty siihen ratkaisuun, että siirtokokonaisuus taltioidaan kokonaisuutena ns. ”media-imagen” arkistojärjestelmään ja järjestelmän sisäisillä sähköisillä eheysmekanismeilla huolehditaan sen muuttumattomuudesta. Tällaisen kokonaisuuden, jolle ei tehdä mitään toimenpiteitä sen muodostuksen jälkeen, säilyttäminen luo selkeän varmistuksen sille, että tiedot säilyvät muuttumattomina ja turvattuina. Tietopalvelun näkökulmasta näiden siirtokokonaisuuksien käsittely olisi erittäin työlästä ja tiedon löytäminen lähes mahdotonta. Sen vuoksi tiedon vastaanottamisen yhteydessä metatiedot ja asiakirjat tallennetaan erilliseen tietokantaan, joka toimii omalta osaltaan eräänlaisena käyttöarkistona.

Sähköisessä muodossa arkistoitavat asiakirjat on siirrettävä arkistolaitokseen säilymisen varmistamiseksi huomattavasti nuorempana kuin paperiasiakirjat nykyisin, jotka siirretään vasta 40-vuoden jälkeen.



Kuva 3. Sähköisen aineiston vastaanottoprosessi

Tämä, ja asiakirjojen sähköinen muoto, asettaa arkistolaitokselle uudenlaisia haasteita tietopalvelun järjestämiseksi. Tavoitteena on tarjota niin arkistoa käyttäville tutkijoille kuin aineistoa muodostaneille viranomaisille sähköinen palvelukanava, jonka kautta aineisto olisi tehokkaasti käytettävissä.

Vuonna 2000 selvitettiin sähköisesti arkistoitavan aineiston määrää. Silloiseen kyselyyn vastasi sähköisten aineistojen osalta 16 % virastoista, jolla oli pysyvässä säilytettävää sähköistä asiakirja-aineistoa yhteensä 4,5Tb. Jos tästä muodostetaan laskennallinen arvio pysyvästi säilytettävästä tietoa-aineistosta kokonaisuutena, niin se olisi ollut noin 25Tb.

Mutta mitä se on vuonna 2010? Osa tästä aineistosta on rekisterimuotoista ja organisaation aktiivisessa käytössä, eikä sitä välttämättä luovuteta, mutta joka tapauksessa tietomäärät ovat jo nyt suuria ja kasvu kiihtyy jatkuvasti.

VAPA-hanke

Sähköistä arkistointi- ja palvelujärjestelmää ei vielä arkistolaitoksessa ole. Järjestelmän suunnitteleminen ja toteuttaminen on perustettu arkistolaitokseen Vapa-hanke, jossa selvitetään sille asetettuja toiminnallisia sekä teknisiä tavoitteita. Hankkeeseen liittyy osana myös yhteistyömahdollisuuksien selvittäminen muiden sähköisiä tietoa-aineistoja arkistovien organisaatioiden kanssa. Vapa-hanke on noteerattu myös

valtion tietoyhteiskuntaohjelmassa ja arkistolaitos on kansallisena toimijana tärkeässä asemassa.

Yhä useampi valtion virasto on siirtymässä nyt yksinomaan sähköiseen asiankäsittelyyn ja arkistointiin. Hankkeen näkökulmasta aikataulu näyttää tämän hetken tietämyksen mukaan siltä, että vastaanottojärjestelmä olisi valmiina käyttöön siinä vaiheessa, kun nämä virastot tekevät ensimmäisiä sähköisiä arkistosiirtoja noin vuonna 2010.

Kirjoittaja työskentelee Kansallisarkistossa sähköisen arkistoinnin ratkaisujen kehittämispäällikkönä. Yhteystiedot (09) 22852 213 tai markus.merenmies@narc.fi.

ESRI Finland

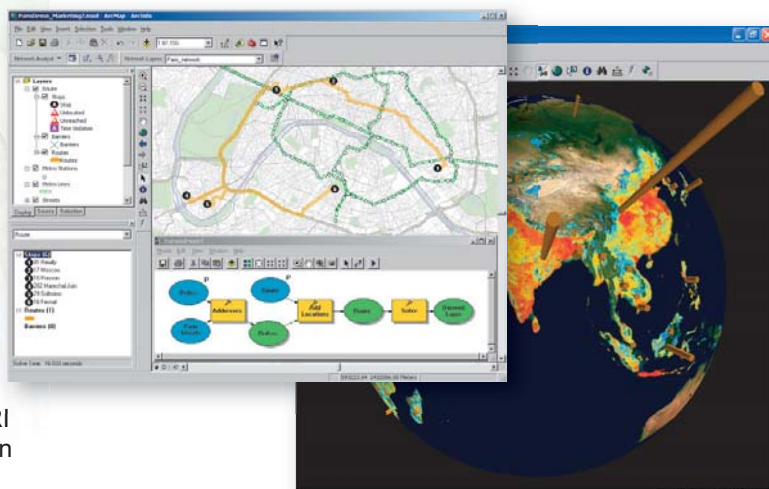
Enemmän irti paikkatiedoista

ESRI:n ratkaisut kokonaisvaltaiseen paikkatiedonhallintaan

- Avoin ja yhteensopiva teknologia
- Kattava tuki eri aineistoformaateille
- Skaalautuu tarpeiden mukaan
- Ratkaisut kaikkiin ympäristöihin ja tarpeisiin

ESRI:n ArcGIS-systeemiarkkitehtuuri sekä Geodatabase-paikkatietokannan ominaisuudet tarjoavat kattavan ja yhteensopivan toiminnallisuuden.

ESRI:n vahva panostus ohjelmistojen kehittämiseen yhdistettynä ESRI Finlandin tuki-, ylläpito- ja koulutuspalveluihin tekevät ESRI:stä varman valinnan



ESRI Finland Oy

Piispanportti 10, 02200 ESPOO
puh. 0207 435 435, myynti@esri-finland.com
www.esri-finland.com, www.esri.com



Musiikin metadata Yleisradion äänilevystössä

Jonne Kulluvaara, Yle

Ajatus musiikin metadatasta on varmaan yhtä vanha kuin ensimmäiset arkistot tai kokoelmat; sisälsivät ne sitten musiikkia käsitteleviä kirjoja, nuotteja tai äänitteitä. Metadata määritellään yleensä tiedoksi, joka kuvailee kohdetta. Musiikin metadatatalla yksilöidään tallenne ja siten helpotetaan sen löytymistä ja käyttöä.

YLEssä Suomen ensimmäinen musiikkiarkisto

Yleisradion vuonna 1934 perustettu äänilevystö on maamme pitkäikäisin musiikkiarkisto, joka on alusta asti itse luonut omat tiedon dokumentointijärjestelmänsä ja -säännöstönsä. Tähän on selkeät syyt: alusta asti hakemistot on tarkoitettu vain yhtiön sisäiseen käyttöön ja palvelemaan ensisijaisesti ohjelmatoimintaa. Tämä on vaikuttanut hakemistojen formaattiin ja tietosisältöön. Tieto on pyritty muotoilemaan käyttäjäsytävällisesti siten, että se on helposti kuuluttajien luettavissa ja käytettävissä edelleen esim. painetuissa lehtien ohjelmatiedoissa.

Yleisradioyhtiöllä on myös monista muista äänitearkistoista ja musiikkikirjastoista poikkeavia velvoitteita mm. ohjelmissa soitetun musiikin raportoinnin suhteen tekijänoikeusjärjestöille (Musa-liittymä). Sekä Teostolla että Gramexilla on ollut omia toivomuksiaan raportoitavista tiedoista ja niiden muotoilusta. Riippumattoman järjestelmän etuja ovat olleet mm. nopea muuntautumiskyky tarpeen mukaan. Esimerkiksi kotimaisten kirjastojen käyttämä kansainvälisen MARC-formaatin suomalaistettu versio, FINMARC, määrää hyvin pitkälle kenttien määrästä ja sisällöstä.

Kaikki Yleisradion äänilevystössä käytettävät tietokannat ovat metatietokantoja: Fono, lainaustark-

kailu ja oman käsikirjaston kirjatietokanta. Äänilevystön kokoelmissa on yli 0,5 miljoonaa äänitettä ja Fono-tietokanta sisältää kaikki äänitteiden luettelointitiedot vuodesta 1974 alkaen, jolloin levystössä siirryttiin mikrokorttipohjaiseen dokumentointijärjestelmään. Lisäksi kantaan syötetään ajoittain myös taannehtivaa materiaalia arkiston alkuajoilta saakka. Tietokannassa on tällä hetkellä yli 1,5 miljoonaa yleistieto-, teos- ja viitedokumenttia. Viimeisin dokumentointijärjestelmä, Windows-pohjainen Fono2-sovellus, toteutettiin yhtiön sisäisenä projektina ja se otettiin tuotantokäyttöön heinäkuussa 1998. Hakujärjestelmänä toimii tanskalainen OpenText Corporationin BRS/Search, jonne dokumentointitiedot päivittyvät useita kertoja vuorokaudessa.

Omista standardeista yhteistyöhön

Äänilevystö on itse luonut omat standardinsa, luokituksensa ja dokumentointisääntönsä, jotka löytyvät työntekijöiden laatimasta Fono-dokumentointioppaasta. Tarpeen mukaan päivitettävä opas on laaja ja yksityiskohtainen ja yhdessä Fono2-dokumentointisovelluksen auktoriteettivalvonnan kanssa mahdollistaa tarkkaan yhtenäistetyn metatiedon tallentamisen. Eritasoista auktoriteettivalvontaa käytetään yhteensä 17:ssä eri kentässä. Kentästä riippuen valvonta on joko pakottavaa, suositavaa tai huomautustyyppistä.

1990-kuvullaäänilevystä alkoi hiljalleen ”avautua” yhtiön ulkopuolelle ja tehdä yhteistyötä musiikkikirjastojen kanssa erilaisissa musiikin luetteloinnin asiantuntijaryhmissä (mm. Valtakunnallinen musiikin luetteloinnin ja sisällönkuvailun Luumu-ryhmä). Käytännössä tämä on johtanut siihen, että soveltuvin osin olemme formoineet metadataamme yhteisten sääntöjen mukaan omiin tarpeisiimme. Äänilevystössä on käytetty lähde- ja tekijätietojen lisäksi Helsingin yliopiston kirjaston julkaisemia YSAa (Yleinen suomalainen asiasanasto) ja Musaa (Musiikin asiasanasto) sekä Suomen musiikkikirjastoyhdistyksen julkaisemia ohjeluetietoja mm. soittimien nimistä ja hankalista musiikkinimistä. Fono-tietokantaan on myös haku- ja katseluoikeudet 7 kirjastolla. Lisäksi tällä hetkellä tutkitaan mahdollisuutta Fonon nykyistä laajempaan käyttöön kirjastoissa yhdessä Helsingin kaupunginkirjaston kanssa.

Levystön metatiedoista vastaavat arkistointilinjan esimies yhdessä kaikkien tiedon dokumentoijien kanssa. Tiedot musiikki-informaattikot tekevät luetteloinnin ohella myös Fono-tietokannan ylläpito- ja haku- ja katseluoikeudet 7 kirjastolla. Lisäksi tällä hetkellä tutkitaan mahdollisuutta Fonon nykyistä laajempaan käyttöön kirjastoissa yhdessä Helsingin kaupunginkirjaston kanssa.

Digiajan myötä yhteiskäyttöön

Lokakuussa 2002 aloitettiin radion digiarkiston tuotantokäyttö. Digiarkisto tarjoaa ohjelmantekijöille mahdollisuuden kuunnella ja käyttää Radion Äänitearkistojen (äänilevystö ja radioarkisto) kokoelmia omilla työasemillaan ja

Koko tietokanta FONO

Dokumentti 1 / 1

Lainauksen tilanne

KART: CD-85862-1:03 Koko äänite
 KEST: 007:03
 FENN: Ei DOKT: Teostiedot KIRJ: 20050818
 KANT: Ei
 OTSI: 'ROUND ABOUT MIDNIGHT
 TEKI: Porter, Cole [1891-1964] (säv).
 TEOS: All of you
 Aluperin elokuvasta Silk stockings.
 DLAA: Ura 1:03 Esikuuntelu Laatuääni
 SMAA: SM7120
 SVUO: SV1954
 LAJI: L5A
 ESIT: Davis, Miles (trumpetti).
 Coltrane, John (tenorisaksofoni). Garland, Red (piano). Chambers,
 Paul (kontrabasso). Jones, Philly Joe (rummut).
 KOKO: KI K25
 TMAA: TM7120
 TVUO: TV1956
 PVUO: PV1957
 RVUO: RV2005
 KAUP: Columbia & Legacy (Sony) / COL 5199572 / (2CD)
 ISRC: USSM15600491
 TIED: Esittelylehtinen.
 HUOM: Arkistokpl:01 käyttökpl:01 kopiokpl:00

Kuva 1. Fono-dokumentti BRS/Search-hakujärjestelmässä.

tuotantojärjestelmillään yhtiön sisäisen verkon kautta kaikkialla Suomessa. Cd-levyiltä valitut urat digitoidaan verkon yli Fono2-soveluksen kautta, jolloin äänitiedostot (MP3-esikuuntelu+ laatuääni) linkittyvät kyseisiin teosdokumentteihin. Laatuäänilatauksen ohessa käyttäjälle tulee XML-dokumentti, joka sisältää ko. teosdokumentin tiedot. Fono-tietokannasta löytyy tällä hetkellä yli 100 000 digitoitua teosta.

Kehittämiskohteita

Hakujärjestelmien kehittyessä yksi kiinnostava kehityssuunta on luonnollisen kielen haku. Siinä tiedonhakijan ei tarvitse enää tietää tiedontallentajien tarkkoja metadatamuotoiluja, vaan asiakas voisi tehdä hakuja normaalilla käyttökielillä. Luonnollisen kielen hakuun liittyvät läheisesti myös käsitteet semanttinen verkko ja itseorganisoiva kartta. Erityisen mielenkiintoisia mahdollisuuksia voisi

syntyä yhdistettäessä useamman YLE:n arkiston tietokannat saman haun piiriin esim. Fono, tv-arkisto, radioarkisto ja valokuva-arkisto.

Muista kehittämiskohteista voisi mainita vielä äänitiedostojen mahdollisen saamisen suoraan levy-yhtiöiden tai jakelijoiden verkon kautta digiarkistoon sekä äänitteiden oheismateriaalin skannaus tietokantaan.

Jonne Kulluvaara työskentelee musiikki-informaattikkona Yleisradion äänilevystössä,
 jonne.kulluvaara@yle.fi



Paikkatieto avain tiedon kokonaisvaltaiseen hyödyntämiseen

Veli-Matti Kiviranta, ESRI Finland Oy

Jopa 80 prosenttiin kaikesta tiedosta liittyy sijainti. Kyseessä voi olla esimerkiksi koordinaatti, katuosoite, postinumero tai kunta. Tällaista sijaintitiedon sisältävää tietoa kutsutaan paikkatiedoksi. Paikkatiedot mahdollistavat tiedon yhdistelemisen, tuottamisen, analysoinnin ja visualisoinnin monipuolisilla menetelmin avaten uusia näkökulmia ja ulottuvuuksia tiedon hallintaan ja hyödyntämiseen.

Paikkatietojärjestelmään voidaan tuoda tietoa jatkojalostettavaksi muista järjestelmistä. Käsitellyt tiedot voidaan edelleen palauttaa käsiteltäviksi esimerkiksi tilasto-ohjelmistoon. Tietoon liittyvä sijainti voi toimia linkkinä esimerkiksi samaan kohteeseen liittyvien eri tietojen välillä. Tietoa voidaan myös hakea, analysoida ja visualisoida sijainnin ja siihen liittyvien ominaisuustietojen perusteella.

Paikkatiedon sujuva käsittely ja hyödyntäminen vaativat avoimen ja yhteensopivan paikkatietojärjestelmän käyttöönottoa. Integroiminen muihin järjestelmiin edellyttää standardinmukaisia rajapintoja sekä tukea eri käyttöjärjestelmille ja aineistomaalle. Tulevaisuudessa paikka-

tieto-toiminnallisuudet integroidaan yhä useammin osaksi organisaatioiden muita tietojärjestelmiä. Tämä takaa tiedon kokonaisvaltaisemman hyödyntämisen. Monilla toimialoilla paikkatiedon merkitys on jo korvaamaton.

Esimerkiksi kaupan alalla paikkatietoja on hyödynnetty jo vuosia monissa liiketoiminnan kannalta kriittisissä toiminnoissa. Myymäläverkon suunnittelu on käytännössä mahdotonta ilman paikkatietoja. Myymälöiden sijointuspäätöksiä tehtäessä on pystytävä varmistamaan, että tietyn ajoikaetäisyyden päässä myymälästä sijaitsee riittävästi ostovoimaa. Toisaalta omien myymälöiden vaikutusalueet eivät saa mennä päällekkäin. Markkinoinnin

tehokas kohdentaminen puolestaan edellyttää, että pystytään paikantaan alueet, joilla asuu haluttuun kohderyhmään kuuluvaa väestöä.

Uudet vaatimukset muuttavat paikkatietoalaa

Paikkatietoalalla on käynnissä rakenteellinen muutos, joka näkyy perinteisten paikkatietojärjestelmien osuuden pienenemisenä ja uusien käyttöalueiden merkityksen lisääntymisenä. Muutos näkyy etenkin järjestelmissä ja niiden arkkitehtuurissa. Työasema-arkkitehtuurista ja hajautuksesta ollaan siirtymässä verkottuneisiin palveluarkkitehtuureihin ja erillisjärjestelmistä siirrytään koko organisaation kattaviin järjestelmiin.

Samalla monimutkaisista asian-
tuntijajärjestelmistä siirrytään
yksinkertaisiin toimintoihin, joissa
monimutkainen toiminnallisuus
piilotetaan helpon ja yksinkertaisen
käyttöliittymän taakse. Paikkatieto
toiminnallisuudet myös integroi-
daan yhä useammin osaksi toimin-
nanohjausjärjestelmiä, asiakkaiden
hallintaa, liiketoiminnan analy-
sointia tai yrityksen logistiikkaa.

Laajamittaisen käytön kynnyksellä

Paikkatietoa hyväksikäyttävien
sovellusten ja palveluiden lisäänty-
minen sekä paikkatietoa hyödyntä-
vien uusien innovaatioiden synty-
minen edellyttää, että eri alojen
ammattilaiset ymmärtävät paikka-
tietojen tarjoamat mahdollisuudet.
Paikkatieto ei enää ole asian-
tuntijoiden erikoistyylikalu, vaan
kiinteä osa eri alojen organisaati-
oiden toimintaa. Tätä edesauttaa
ohjelmistotoimittajien tarjonnan
kehittyminen. Samat paikkatieto-
toiminnot on saatavilla usein sekä
valmiissa tuotteissa että erillisinä
komponentteina, jotka voidaan
upottaa osaksi asiakkaan tarvit-
semaa toiminnallisuutta.

Erilaiset ilmaiskarttapalvelut,
kuten Map Quest ja Google
Earth, ovat osaltaan vieneet alaa
eteenpäin. Tietoisuus paikkatie-
tojen mahdollisuuksista kasvaa.
Samalla kasvavat myös vaatimukset
koskien erilaisten palveluiden
toiminnallisuutta. Organisaatiot
alkavat tunnistaa toiminnastaan
paikkatietotarpeita, joihin tarvi-
taan innovatiivisia ratkaisumal-
leja. Paikkatietojen hyväksikäyttö
mahdollistaa siis organisaatioiden

kehittää toimintaansa uudesta
näkökulmasta.

Taustalla ovat usein paineet
toiminnan tehostamiseksi ja
resurssien käytön optimoimiseksi.
Investoinneilta vaaditaan entistä
enemmän tuottoa. Tietoa tuote-
taan enemmän kuin koskaan,
mutta siitä ei saada otettua kaikkea
hyötyä irti. Merkittävä osa kaikesta
tiedosta voidaan määritellä paikka-
tiedoksi, jolloin sen kokonaisval-
tainen hyödyntäminen edellyttää
paikkatietotoiminnallisuuksien
soveltamista.

Paikkatietojen merkitys kasvaa

Paikkatiedon integroitua entistä
tiiviimmin osaksi organisaati-
oiden tietojärjestelmiä nousevat
niille asetetut vaatimukset uudelle
tasolle. Vielä joitakin vuosia sitten
organisaation laajuinen paikkatie-
tojärjestelmä saatettiin toteuttaa
hankkimalla tarvittavat työase-
malisenssit paikkatietoja työssään
tarvitseville.

Tulevaisuudessa tullaan siirtymään
keskitettyyn paikkatiedonhallin-
taan ja paikkatietopalveluihin.
Tarve erillisille työasemaohjelmi-
sille poistuu, mikä näkyy muun
muassa alhaisempina ylläpitomak-
suina. Paikkatietotoiminnallisuuk-
sia tullaan käyttämään upotettuna
osana muita sovelluksia tai esimer-
kiksi Internet-selaimen kautta. Eri
käyttäjärhyhmille voidaan määri-
tellä yksilölliset käyttöoikeudet ja
räätyä löydyt valmiin toiminnot ja
työkalut. Hyödyntääkseen paik-
katietoja ei enää tarvitse olla alan
asiantuntija.

Sulautuessaan muihin tietojär-
jestelmiin paikkatieto muodostaa
toiminnan kehittämisen alustan.
Paikkatietotoiminnallisuudet
tulevat kiinteäksi osaksi organi-
saation prosesseja kaikille tasoille,
aina tiedon tuottamisesta ja analy-
soinnista sen hyödyntämiseen ja
jakeluun, niin asiakaspalvelussa,
toiminnan suunnittelussa kuin
päätoimintonsa tuessakin.

Aineistot, osaajat ja ideat ratkaisevat

Paikkatietopalveluiden kehittä-
misen esteenä on perinteisesti ollut
sopivien aineistojen saatavuus ja
hinta. Tämä on vaikeuttanut uusien
paikkatietopohjaisten palveluiden
kehittämistä ja jopa estänyt uusien
innovaatioiden tuotteistamisen.

Viime vuosina kehitys on mennyt
järkevään suuntaan. Yhä katta-
vampia aineistoja on saatavilla
useilta eri toimittajilta, mikä näkyy
ostajan kannalta suotuisana hinta-
kehityksenä. Edulliset paikkatie-
toaineistot mahdollistavat uusien
tuotteiden ja palvelukonseption
kehittämisen sekä uusien yritysten
markkinoille tulon. Tämä johtaa
paikkatietoratkaisuiden ja -palve-
luiden yleistymiseen. Samalla alan
koulutuksen saaneiden osaajien
sekä erilaisten tutkimus- ja tuote-
kehityshankkeiden lisääntyminen
edesauttaa uusien innovaatioiden
syntymistä.

Yritysten ja yhteiskunnan tuot-
tamien palveluiden on pystyttävä
vastaamaan muuttuviin tarpeisiin.
Toimintoja ulkoistetaan ja organi-
saatioiden sisällä siirrytään tilaaja-
toimittaja-malliin. Lisäksi perin-
teisestä asiakaspalvelusta ollaan

siirtymässä itsepalvelukulttuuriin, mikä näkyy palvelujen siirtymisenä Internetiin.

Asiakaspalvelun ulkoistaminen on jo arkipäivää monissa yrityksissä. Tämä merkitsee usein sitä, että koko maan asiakkaita palvelaan keskitetysti yhdestä paikasta. Asiakaspalvelun työntekijöillä ei välttämättä enää ole työn kannalta tärkeää paikallistuntemusta. Tämän kaltaisissa tilanteissa paikkatietojen rooli nousee korvaamattomaksi.

Metadata takaa paikkatiedon käytettävyyden

Metadata dokumentoi tiedot mm. aineistojen sisällöstä, laadusta ja tyypistä. Erilaisten paikkatietoaineistojen määrän voimakas kasvu ja tiukentuneet vaatimuksen tiedon sujuvan jakamisen järjestämiseksi edellyttävät paikkatietojärjestelmiltä kehittynyttä teknologiaa metatiedon hyödyntämiseksi.

Metatieto on tärkeässä asemassa arvioitaessa paikkatiedon käytettävyyttä. Sen avulla esimerkiksi aineistojen dokumentointi ja sopivien aineistojen löytäminen nopeutuu. Tukemalla metadataa tiedontuottajat voivat julkaista tietoa aineistoista ja käyttäjät löytävät tarvitsemansa aineistot.

Paikkatietoa hyödyntävien organisaatioiden on tiedettävä täyttääkö paikkatietoaineisto tarpeet ja laadulliset vähimmäisvaatimukset. Näin metatiedon merkitys on suuri osana päätöksenteon tukea. Metatiedon julkaiseminen helpottaa myös aineistojen jakamista organisaatioiden välillä.

Tiedonhallinnan näkökulmasta metatieto on välttämätöntä ylläpitäessä organisaation paikkatietoinvestointeja.

Metatieto on merkittävässä asemassa osana jokaisen paikkatietojärjestelmän implementointia. Tämän vuoksi sen tulisi olla standardi osa jokaista paikkatietojärjestelmää eikä vain erikseen tarjottava lisätoiminnallisuus. Työkalut metatiedon luomiseen ja hallintaan ovat avainteknologioita paikkatiedon sujuvaan jakamiseen. Se mahdollistaa mm. käyttöoikeuksien määrittämisen erityyppisiin aineistoihin.

Uusi ajattelutapa, uudet vaatimukset

Teknologian kehittyminen, tiedon määrän kasvu ja kasvavat vaatimukset heijastuvat paikkatietoalan murroksena. Teknologian ja sen sovellusmahdollisuuksien tullessa tutuksi yhä laajemmalle yleisölle syntyy uusia innovaatioita vastaamaan muuttuviin tarpeisiin. Samalla paikkatieto tullaan näkemään kokonaisuuden osana.

Tarvitaan entistä enemmän sovelluksia ja palveluita, jotka pohjautuvat ajantasaiseen tietoon. Palvelut on pystyttävä upottamaan olemassa oleviin sovelluksiin ja monimutkainen toiminnallisuus on pystyttävä toteuttamaan yksinkertaisen käyttöliittymän avulla. Palvelut on saatava helposti ja kohtuullisilla kustannuksilla suurenkin käyttäjäjoukon saataville.

Vastatakseen tarpeisiin paikkatietojärjestelmätoimittajilta vaaditaan kustannustehokkaita ja yhteen-

sopivia järjestelmiä, jotka tukevat asiakkaiden käytössä olevia laitteistoja, käyttöjärjestelmiä sekä aineistoja. Avoin ja standardeihin perustuva teknologia edesauttaa ja helpottaa järjestelmien integrointia ja takaa investointien säilymisen myös tulevaisuudessa. Lähtökohtana on oltava, että organisaation paikkatietoperusta mukautuu ja skaalautuu tarpeiden mukaan.

ESRI Finland Oy toimii ESRI:n paikkatietoratkaisuiden edustajana ja jälleenmyyjänä Suomessa. Yhtiö tarjoaa ESRI:n teknologiaan perustuvia paikkatietoratkaisuja kattavasti eri toimialojen tarpeisiin. ESRI Finlandin ydinliiketoimintaan kuuluvat lisäksi kattavat koulutus-, konsultointi-, tuki- ja ylläpitopalvelut. ESRI Finland toimii tiiviissä yhteistyössä suomalaisten liiketoimintakumppaneittensa kanssa, joita ovat mm. WM-data, TietoEnator ja Fifth Element. Yhteistyökumppaneihin kuuluu myös tietotekniikan suurimpia laite- ja ohjelmistotoimittajia. Lisätietoa: www-esri-finland.com

ESRI (Environmental Systems Research Institute, Inc.) on maailman johtava paikkatietoratkaisuiden toimittaja. ESRI tarjoaa ratkaisut kaikkeen paikkatiedonhallintaan, aina yhden käyttäjän työasemasovelluksista koko organisaation kattaviin Enterprise-GIS -järjestelmiin. ESRIllä on takanaan yli 30 vuoden kokemus, sen tuotteita on käytössä yli 300 000 organisaatiossa maailmalla ja jälleenmyyjä yli sadassa maassa. Lisätietoja ESRIstä: www.esri.com

Tietämyksen mallintaminen semanttisessa webissä

Tomi Kauppinen, Tuukka Ruotsalo, Mirva Salminen, TKK Viestintätekniikka ja Helsingin Yliopisto

Semanttisessa webissä mallinnetaan semantiikkaa koneiden käyttöön toisin kuin perinteisessä webissä, jonka dokumentit ovat ihmisille tarkoitettuja. Näin koneet saadaan valjastettua sekä tietämyksen tarkkaan tallennukseen että valtavien tietomäärien käsittelyyn. Tässä käytetään apuna ontologioita eli koneymmärrettävästi ilmaistuja käsitteitä ja niiden välisiä suhteita. Sovelluskohteina ovat esimerkiksi tietämyksen ja aineistojen hallinta sekä elektroninen kaupankäynti.

Semanttisen webin teknologiat ja ontologiat

Semanttisessa webissä [1] mallinnetaan semantiikkaa koneiden käyttöön toisin kuin perinteisessä webissä, jonka dokumentit ovat ihmisille tarkoitettuja. Semantiikka mallinnetaan ontologioilla, eli koneymmärrettävästi ilmaistulla käsitteillä ja niiden välisillä suhteilla.

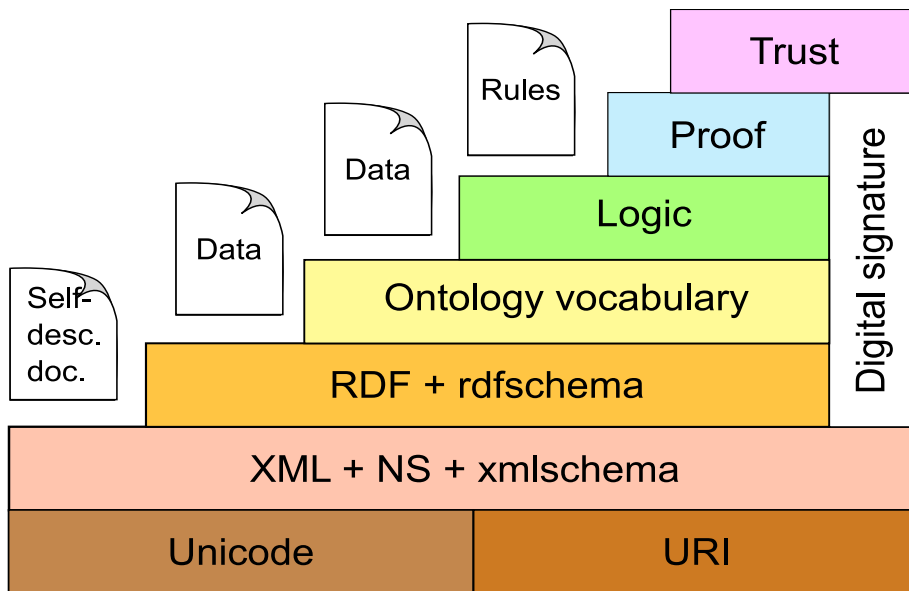
W3C:n johtajan Ivan Hermanin mukaan [2] semanttiselle webille on tapahtumassa sama kuin webille tapahtui useita vuosia sitten: kunnollisten työkalujen, järjestelmien ja sovellusten myötä liiketoiminta vakiintuu. Tämä on mahdollista, koska semanttisen webin teknologiat ovat jo tulossa yritysten perusteknologioiksi.

Semanttisen webin teknologiat [1] helpottavat ontologiatyön tekoa käytännön tasolla, toisin sanoen ne tarjoavat välineet, joiden avulla on mahdollista määrittellä semanttisen tiedon eli käsitteet sekä niiden väliset suhteet täsmällisesti ja koneymmärrettävästi. Käsitteet identifioidaan yksiselitteisesti URIen (Universal Resource Identifier) avulla, jolloin niihin voidaan viitata globaalisti. Ontologioiden kuvaamiseen on kehitetty esimerkiksi OWL (Web Ontology Language).

Ontologian käsitteitä voidaan käyttää esimerkiksi metatiedon sisällönkuvailussa aina julkisten museoiden ja kirjastojen kokoelmista yksityisten ihmisten musiikkikokoelmiin ja yritysten verkossa

tarjoamiin palveluihin asti. Kun mahdollisimman moni jakaa saman käsitteistön, jossa käsitteillä on identifioivat tunnisteet, voidaan järjestelmien integrointi suorittaa tehokkaammin. Lisäksi voidaan saada selville aineistojen välisiä mielenkiintoisia ja yllättäviäkin suhteita päättelykoneiden avulla.

Kuvassa 1 on esitetty Tim Berners-Leen esittelemä Semanttisen webin ”teknologiakakku” (layer cake) [3]. Kakku kuvaa kuinka perusteknologioiden Unicoden ja URIn varaan on rakennettu syntaksikerros XML: n nimiavaruuksien (NS) avulla ja tietomallikerros RDF:n ja RDF skeeman avulla. Tällä hetkellä rakennetaan erityisesti sekä seuraavina tulevaa ontologiakerrosta



Kuva 1. Semanttisen webin teknologiakakku (layer cake) Tim Berners-Leen mukaan [3].

(ontology vocabulary) että logiikka-kerrosta (logic). Seuraavaksi rakennetaan todistuserrosta (proof) ja lopulta ideana on rakentaa luottamuskerros (trust).

Ontologioilla mallinnetaan siis käsitteitä ja niiden välisiä suhteita. Esimerkiksi purjehduksessa käytetään purjevenettä, jonka yksi osa on purje. Purjevene puolestaan on eräänlainen vene. Samoin moottoriveneen osa on moottori ja se myös on vene. Suhteista ja käsitteistä muodostuu mallinnuksen tuloksena verkko, johon perustuen voidaan tehdä päätelmiä käyttäen logiikkaa ja heuristisia sääntöjä. Päätelykone voi tehdä ontologian avulla eron esimerkiksi purjeveneen ja moottoriveneen välillä sekä toisaalta nähdä missä mielessä niillä on yhteisiä piirteitä. Ontologioiden käyttö mahdollistaa myös vaikeiden ja pitkien päätelyketjujen teon: esimerkiksi kuvaa laajemman ontologian perusteella voitaisiin löytää yhteys navigointivälineiden ja purjeveneen välillä.

Yläontologiat yhdistävät kohdealue tietämystä

Yksittäisten sovellusten käyttämä tieto esitetään usein kohdealueella kuvaavalla ontologialla tai tietomallilla. Semanttisen webin keskeisenä ajatuksena on yhdistää metatietoja verkossa, eli kohdealueiden tietomalleja semanttiseksi verkoksi, jossa päättely ja automaattinen tietojen yhteiskäyttö on mahdollista.

Eräänä ratkaisukeinona kohdealueiden yhteentoimivuuden saavuttamiseksi on tarjottu ylä-ontologioita. Ne tarjoavat yhteisen, laajennettavan peruskäsitteistön, joiden avulla mahdollistetaan erilaisten sovellusten ja niiden hyödyntämien tietomallien yhteentoimivuus [4]. Ideana on siis, että yläontologiassa esitellään joukko käsitteitä ja suhteita, joita tarkennetaan kohdealuekohtaisissa ontologioissa. On epärealistista ajatella, että kaikki jonkin kohdealueen tietomallit voidaan mallintaa yhdessä sovitulla tavalla. Tämän vuoksi tarvitaan

keinoja kytkeä uusia käsitteitä ja kohdealueontologioita yhteen.

Esimerkiksi yläontologian käsitteen vene, suhteen voitaisiin toisessa kohdealueessa kuvata tarkemmin käsitteet purjevene ja moottorivene. Käyttämällä sopivaa sääntökieltä voidaan luoda myös monimutkaisempia käsittemalleja. Esimerkiksi käsite purjehtija voitadaan kuvata käsitteen purjehtiminen suhteen säännöllä, jossa määrätään purjehtija käsitteen ilmentymiksi kaikki ne ihminen-käsitteen ilmentymät, jotka on kytkeyty tapahtumaan purjehtiminen millä tahansa kaarella. Näin mahdollistetaan toisistaan poikkeavien kohdealueontologioiden käyttäminen osana laajempaa tietämyskantaa. Menetelmien kehittäminen voidaan lisäksi kohdistaa ontologioihin kohdealuekohtaisten tietomallien sijaan.

Esimerkkinä kohdealueontologioista voi mainita, että tiedonhallinnan alueella on DAMA Finlandissa käynnistetty tiedonhallinnan keskeisen käsitteistön kokoaminen ontologiaksi.

Muuttuvan maailman mallintaminen ontologioilla

Ontologioilla pyritään kuvaamaan maailmaa --- mutta ongelmaksi muodostuu helposti se, että mallinnettava maailma muuttuu. Esimerkissämme vene-käsitteen tarkentavina käsitteinä voi nykyisin olla käsitteet moottorivene ja purjevene, mutta aikaa ennen moottoriveneen keksimistä kuvaavassa ontologia-aikasiivussa käsite moottorivene puuttuu.

Myös maantieteellinen käsitteistö muuttuu. Jos esimerkiksi lähitulevaisuudessa tehdään päätös joidenkin kuntien yhdistämisestä ja siten kuntien kokonaismäärän vähentämisestä, niin myös paikkojen ontologiaa on muokattava. Tämänkaltaisten muutosten mallinnusta voidaan tehdä ontologia-aikasarjan ja muutossiltojen [5] avulla ja hyödyntää tätä tietoa hakukoneissa. Niiden avulla on esimerkiksi saatu selville, että nykyinen Lappeenranta peittää 1900-luvun vaihteen Viipurista n. 12.5%. Sen ajan Viipurilla indeksoitu museoesineistö on siis todennäköisesti osittain peräisin nykyisen Suomen alueelta --- tieto, jota voidaan hyödyntää aineistojen hallinnassa. Päättelykoneita käyttämällä voidaan semanttisen webin ontologioista siis kaivaa yllättävää tietämystä ja tarkentaa siten hakutuloksia.

Lisätietoja:

W3C: <http://www.w3c.org/2001/sw/>

Semanttisen laskennan tutkimusryhmä (SeCo): <http://www.seco.tkk.fi/>

FinnONTO-hanke: <http://www.cs.helsinki.fi/group/seco/ontologies/>

Dama Finland: <http://www.dama-finland.org/>

Lähteet:

[1] Tim Berners-Lee, Jim Hendler, and Ora Lassila. The Semantic Web. Scientific American, 284(5):34–43, May, 2001.

[2] Ivan Herman. Towards the Semantic Web. Invited talk and tutorial at Web Intelligence symposium. The Finnish AI Conference, 20 year Anniversary (STeP 2004), Heureka - the Finnish Science Centre, Vantaa, Finland, September 2-3, 2004.

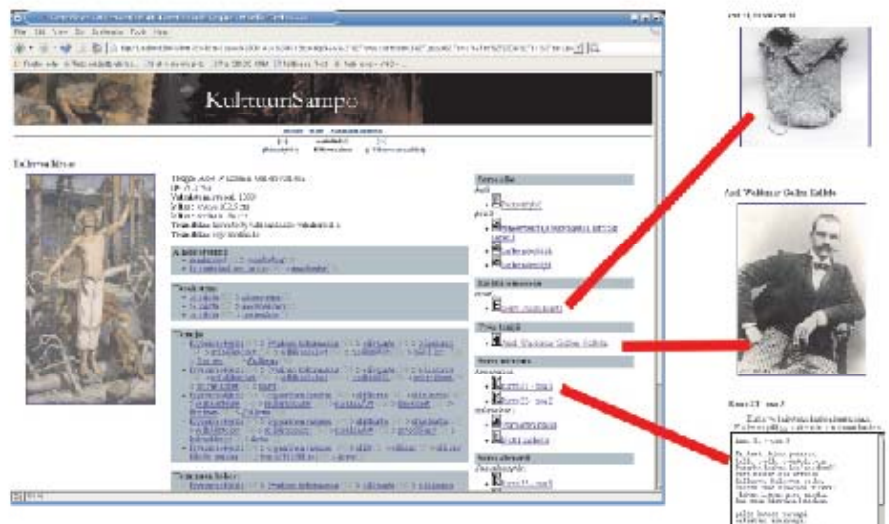
[3] Tim Berners-Lee. Semantic Web Wedding Cake (or "layer cake") in a conference talk, XML 2000 conference, Washington DC, 2000.

[4] Eero Hyvönen, Arttu Valo, Ville Komulainen, Katri Seppälä, Tomi Kauppinen, Tuukka Ruotsalo, Mirva Salminen and Anu Ylisalmi: Finnish National Ontologies for the Semantic Web - Towards a Content and Service Infrastructure. Proceedings of International Conference on

Dublin Core and Metadata Applications (DC 2005), Madrid, Spain, Sep, 2005.

[5] Tomi Kauppinen and Eero Hyvönen: Modeling and Reasoning about Changes in Ontology Time Series. Chapter in book. Ontologies in the Context of Information Systems (Rajiv Kishore, Ram Ramesh and Raj Sharman (eds.)), Springer-Verlag, Berlin, Dec, 2005. In press.

Tutkijat Tomi Kauppinen, Tuukka Ruotsalo ja Mirva Salminen työskentelevät Teknillisen korkeakoulun viestintätekniikan laboratoriossa ja Helsingin yliopiston tietojenkäsittelytieteen laitoksella toimivassa semanttisen laskennan tutkimusryhmässä (SeCo). Tomi Kauppinen on toiminut DAMA Finlandin hallituksessa sen perustamisesta lokakuusta 2004 lähtien ja XML Finland ry:n hallituksessa vuosina 2000-2005.



Kuva 2. Esimerkki Suomalaisesta ontologiasta.

KulturiSampo linkittää automaattisesti kulttuurisisältöjä. Kuvassa Ateneumin kokoelmiin kuuluva maalaus Kullervon kirous on yhdistetty Kansallismuseossa olevaan tuohikonttiin (jollainen esiintyy maalauksessakin), Akseli Gallen-Kallelan biografiaan tietoihin ja ontologisesti kuvattuun, Kalevalassa olevaan runoon Kullervon kirouksesta.

Eero Hyvönen



Tiedonhallinnan Ontologiaa tekemässä

Leyla Akgez, Nokia

DAMA Finlandin jäsenistö on innostunut kokeilemaan ontologioiden tekoa käytännössä, koska se tuntuu hyvältä tavalta luoda hajautetusti yhteistä ymmärrystä tietyn aihealueen käsitteistöstä. Tässä artikkelissa kuvaan prosessia, jolla olemme lähteneet liikkeelle.

Miten ajatus Tiedonhallinnan Ontologiasta heräsi?

DAMA Finlandin tarkoituksena on yhdistää tiedonhallinnan ammattilaisia ja tiedonhallinnasta kiinnostuneita. Mutta jo DAMA Finlandin perustamisesta lähtien oli selvää, että jollain tavalla olisi hyvä saavuttaa jäsenistön kanssa yhteisymmärrys siitä mitä tiedonhallinnalla tarkoitetaan. Eri yrityksissä ja yhteisöissä ”tiedonhallinnan” alle niputetaan toisistaan eroavia asioita, tekemistä ja tekijöitä. Pohdimme ensin, että kenties järjestämme jokaisesta tärkeäksi kokemastamme alueesta koulutusession jossa käytäisiin läpi perusasioita tuosta alueesta. Kysymykseksi nousi pian se, että kuka määrittää mikä alue on tärkeä ja mitä siihen alueeseen kuuluu.

Yhdistyksen hallitus ei sitä yksin voi tehdä, ja perinteiset jäsenkyselyt eivät yleensä tuota kovin informatiivista tulosta. Jollain tavoin jäsenistö oli saatava mukaan toiminnan määrittelyyn.

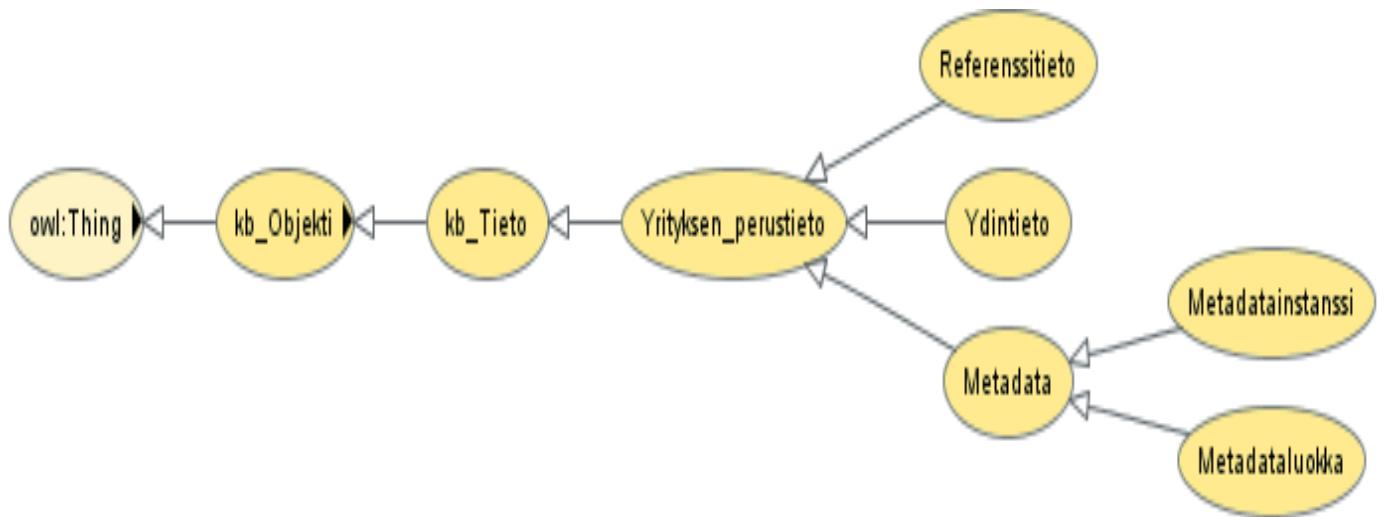
Toisaalta DAMA Finland haluaa tehdä myös yhteiskunnallisesti merkityksellistä työtä, mitä ajatellen hallitus pohti ankarasti miten voisimme auttaa suomalaisia yrityksiä ja yhteisöjä tiedonhallinnan alueella. Totesimme että jos meillä oli vaikeuksia määritellä mitä tiedonhallinta on ja mitä asioita alueella liikkuu, tämä voi olla monille yrityksille ja yhteisöille myös se ensimmäinen huolenaihe. Jos tämän onnistuisi jotenkin järkevästi kuvaamaan, ja sitten jakamaan kaikille suomalaisille tiedon

tarvitsijoille niin siitä olisi varmasti hyötyä.

Tässä vaiheessa idea jonkinlaisesta laajasti jaetusta termityöstä oli selkeä, ja sitten kun tutustuimme ontologia-tekniikoihin, niin oli selvää että haluaisimme kokeilla niitä tässä. Ensimmäisen kerran ontologioihin perustuvat hankkeet olivat esillä jo perustamiskokouksessamme, siellä Eero Hyvönen esitteli Museo Suomi-hanketta ja sen taustalla olevia ontologioita.

Tuumasta toimeen

Keväällä järjestettiin jäsenilta 26.5. RDF-tietomallinnuksesta OP-ryhmän tiloissa, jossa Janne Saarela ja Tomi Kauppinen kertoivat RDF-mallinnuksen teoriaa ja käytännön sovellutuksia, Janne yritysmaa-



Kuva 1. Ontologioissa käsite määärätty sen suhteiden kautta.

ilman puolelta ja Tomi yliopistojen projekteissa. Paikalla oli paljon jäseniä, ja monet sanoivat että haluaisivat päästä itse kokeilemaan näiden teknologioiden käyttöä.

Torstaina 2.6. oltiinkin jo tositoimissa, kokoonnuimme Nokian tiloihin ja monilla oli tietokone mukana, jonne kutsun ohjeiden mukaan oli asennettu Protégé – työkalu. Tomi Kauppinen, Ville-Pekka Komulainen ja Tuukka Ruotsalo esittelivät lyhyesti työkalua sekä HIITin (The Helsinki Institute for Information Technology) ontologia-aiheisia projekteja, jotka saattaisivat olla meidän kannalta kiinnostavia. YSO-projektissa (Yleinen Suomalainen Ontologia) luodaan yleiseen suomalaiseen asiasanas- toon osittain perustuva ”katto-ontologia” johon tällaisia tietyn aihealueen tarkempia ontologioita voisi ”ripustaa”. Ajattelimme heti että tämä Tiedonhallinnan Ontologia olisi yksi sellainen ripustettava ontologia, ja käytäisimme muutamia YSO:n peruskäsitteitä (esim. Toiminta, Toimija jne.) pohjana, vaikka YSO ei vielä

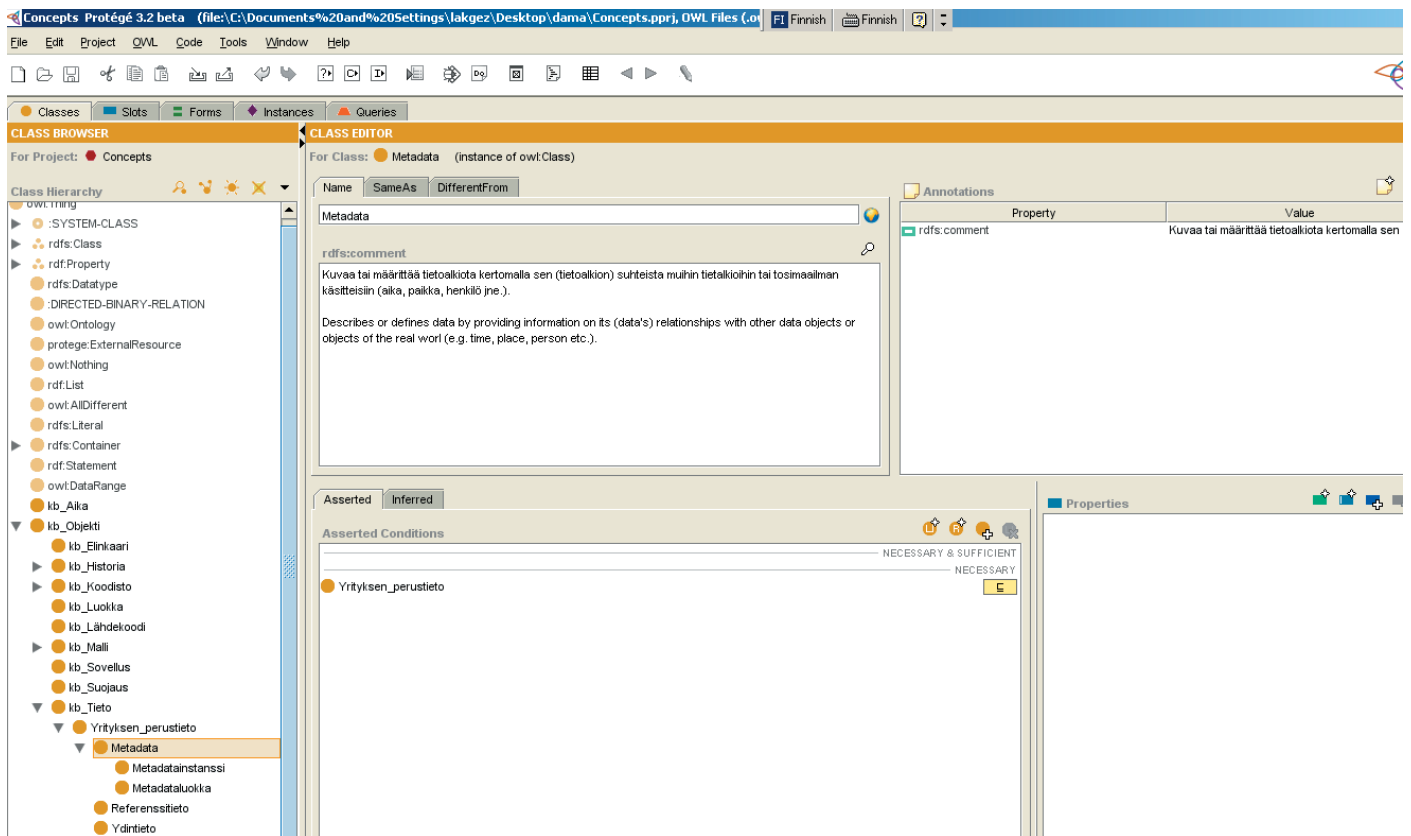
olekaan vielä julkaistu. Toinen esitelty projekti oli OnKi-projekti, jossa luodaan teknologiaa jolla voisi julkaista oman ontologiansa muiden käyttöön. Tämäkin olisi meille hyödyllistä, tavoitteenahan meillä on julkaista ontologiamme kaikkien suomalaisten yritysten ja yhteisöjen käyttöön. Loppuaika työryhmässä käytettiin tiedonhallinnan käsitteistä keskustellen.

Kesän aikana lomailtiin, mutta lomailevatkin mielet pohtivat tärkeitä kysymyksiä ontologioista. Hallituksen puheenjohtaja Pekka Valta oli innostunut opiskelemaan Protégén käyttöä ja tehnyt oman pikkuontologian työssä kohtaamistaan käsitteistä. Samaan aikaan hallitus mietti teknologia- valintaansa. Protégé työkaluna on sinänsä hyvä, että se on kaikkien saatavilla ja hyvin ohjeistettu, mutta senkin käytössä saattoi valita erilaisten tapojen joukosta, mm. käyttäkö Protégén omaa formaattia, RDFää vai OWLia. Ei-standardina tallennusmuotona tuo työkalun oma muoto oli pois laskuista ja RDFn ja OWLin välinen keskustelu kääntyi OWLin

voitoksi. Periaattessahan OWL sisältää kaikki RDFn mahdollisuudet, mutta laajentaa niitä eteenpäin.

Käytännön työskentelytavat

Syksyn ensimmäinen ontologia-työpaja venyi ajateltua myöhemmäksi logististen vaikeuksien takia, mutta osanotto oli vieläkin innostunutta ja keskustelu vilkasta. Täytyy sanoa yhdistyksen jäsenten kunniaksi, että kaikki työpajassa olleet olivat innostuneita kokeilemaan tätä uutta teknologiaa ja osallistumaan ihan raakaan työhönkin. Kertasimme mihin asti olimme päässeet (eli hyvin varhaiseen alkuun), ja sovimme miten jatkaisimme käytännössä. Ontologioissa käsite määärätty sen suhteiden kautta, joten käsitteitä voidaan määritellä hyvinkin hajautetusti. Jos jollain kahdella käsitteellä on yksikin sama suhde, ne voidaan sen kautta ”kytkeä” toisiinsa. Sovimme että Tiedonhallinnan Ontologia syntyisi iteroiden jäsenistön yhteistyönä. Koska mitään ”oikeapista” tapaa kehittää ontologioita ei oikeastaan ole määritetty, ajattelimme kokeilla menetelmää, jossa



Kuva 2. Ruudunkaappaus Protégéstä, tältä mallintaminen näyttää.

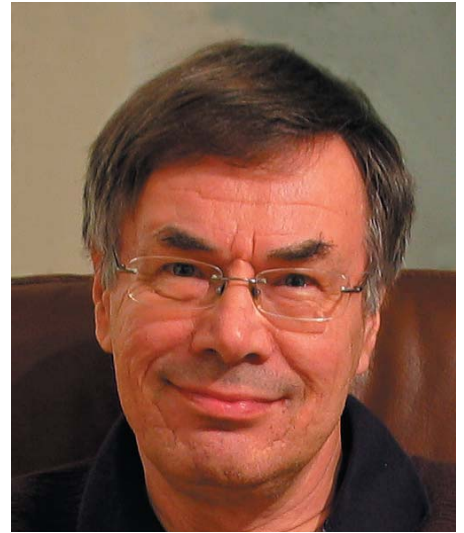
jokainen ottaa itselleen läheisen käsitteen (esimerkiksi alueelta jossa työskentelee) ja pohtii mitä Toimintaa, Toimijoita, ja Asioita aiheeseen liittyy, ja sitten katsomme kuinka paljon joukossa on yhdistäviä suhteita. Sovimme että kaikki lähettävät määräpäivän sen hetkisen tilanteensa, vaikkakin keskeneräisen, koska ontologioissa ei oikein ole edes tarkoitus tyhjentävästi selittää asiaa, vain sen verran että käsitteen pystyy tunnistamaan. Teknisesti yhdistäminen onnistuu Protégé-työkalulla helposti, ja yhdistämisen jälkeen on helpompi nähdä aukkoja ja tärkeitä solmukohtia ontologiassa. Ontologia julkaistaan DAMA Finlandin verkkosivuilla ja sitten istutaan pöydän ääreen keskustelemaan siitä ja jakamaan seuraavia tehtäviä.

Innokkaat mukaan tekemään

Tähän työhön voi kuka tahansa osallistua, ja tämä on hyvä tapa oppia käytännössä ontologioiden tekemistä ja hajautetun työskentelyn metodologiaa. Tiedonhallinnan ammattilaiselle tällainen ontologiatyö auttaa myös oman työn selventämistä, ja tukee hyvin päivittäistä työtäkin. Sen huomasin itse kun ryhdyin kuvaamaan omaa työskäntäni, metadatan käsitettä ontologian keinoin. Ohessa on kuva tämän hetkisestä pohdinnasta. Kuten aiemmin painotin, ontologioiden maailmassa ei ole oikeita tai väriä määritelmiä, idea on demokraattisesti antaa jokaisen määritellä asia omasta näkökulmastaan, ja niistä suosituimmat sitten jäävät elämään ja kenties leviävät laajempaankin käyttöön. Lisäksi pitää korostaa, että se on vain työstöversio, joka ei ole edes oma valmis näkemykseni

aiheesta, vaan on tässä esimerkinomaisesti kuvaamassa työskentelytapaa.

Kirjoittaja Leyla Akgez on FM ja DAMA Finlandin hallituksen jäsen. Leyla Akgez työskentelee tällä hetkellä Nokialla metadata-asiantuntijana. Lisätietoa ontologiatyöstä löytyy DAMA Finlandin kotisivuilta (<http://www.damafinland.org/Ontologia.html>) ja mukaan voi ilmoittautua esimerkiksi Leylalle, leyla.akgez@nokia.com.



Ensikokemuksia ontologioista

Pekka Valta, Dama Finland

Tuoreena ammattiyhteisönä Dama Finland on todennut, että pitäisi saada aikaiseksi tiedonhallinnan käsitteistö ja vieläpä Ontologian muotoisena. Mistä oikein on kysymys ja miksi innostua moiseen puuhaan?

Työni puolesta käytän hyvin aktiivisesti nettiä tiedon hankintaan ja tuska informaation määrästä ja sirpaleisuudesta on suuri. Miten suodattaa pois vanhentunut ja kloonitieto? Miten löytää omaperäiset ja uutta ennakoivat tiedon lähteet? Onko jotain pysyvää, joka kestää hypeilyn ja kvartaalitalouden puhurit? Olen vakuuttunut, että vastaus on Ontologia.

Uusi juttu?

Ontologia palaa käsitteenä peräti vuoteen 384 e.a.a ja kreikkalaiseen Stragiruksen siirtokuntaan? Nyt valistuneilla lukijoilla sytytti, kysehän on Himasta ja Gatetsiä suuremman ajattelijan Aristoteleen syntymätiedoista. Hän pohti jo systemaattisesti asioiden olemaisuutta ja niiden keskinäisiä suhteita. Myöhemmät filosofit työstivät ontologiaa objektien,

niiden ominaisuuksien ja keskinäisten suhteiden rakenteiden tieteenä. Viime vuosien aikana ontologia on saanut tekniikkana käyttöä laki-, lääke-, insinööri- ja taloustieteiden alueilla. Ja tietenkin tietojenkäsittelyn alueella, DAMA yhtenä aktivistina.

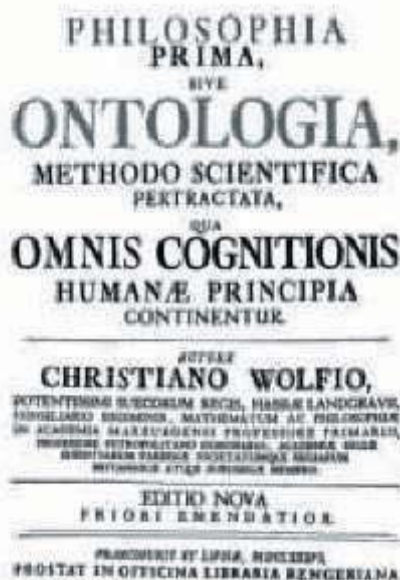
Tietojenkäsittelylle on tyypillistä tekniikoiden, käsitteiden ja lyhenteiden nopea lisääntyminen ja vanhentuminen. Tämä tuo mielenkiintoisia erityispiirteitä DAMAnkin käsitetyöhön. Pitäisi löytää pysyviä rakenteita ja välttää yksityiskohtien näpertely. Esimerkiksi Tietotekniikan liitto ry:n sanastotoimikunnan laatimaa ATK-sanakirjaa on päivitetty jatkuvasti vuodesta 1966 lähtien ja samainen ATK-sanakirja sisältää jo lähes 5000 tietotekniikan termiä ja lyhennettä englanniksi ja suomeksi

sekä noin 4000 vastinetta kuudella muulla kielellä, yhteensä siis yli 30 000 termiä.

Kokeiluja

Olen kesästä lähtien perehtynyt ontologioihin käytännön kokeilujen kautta. Työkaluna on käyttänyt DAMAn hyödyntämää Protege-ohjelmaa, sen ohjeet löytyvät DAMAn kotisivuilta. Kokeiluluonteisia suppeita ontologioita olen rakennellut ns. "REA-viitekehystä", "verkkosivustosta" ja "yhdistyksestä". Oma työskentely on nimittä'in antoisampaa, jos aloitti itselleen hyvin tutusta aihealueesta ja opettelee työkalunsa (Protegen) ominaisuudet riittävän varmasti.

Alkuhapuilun jälkeen totesin viisaaksi opiskella kunnolla semanttisen webin RDF-stan-



Kuva 1. Christian Wolff's *Philosophia prima sive Ontologia* (1729) – Kirjan otsikossa on ensi kertaa käsite "ontologia".

dardin perusideat ja niiden jatko-kehitykset, lähinnä sitä ilmaisuvoimaisempi OWL-notaation.

Pian selvisi myös, että ontologioiden pysyvä arvo ei ole em. ATK-sanakirjan tapaisissa sanastoissa eikä laajoissa verbaalisissa käsitteiden kuvauksissa, vaan loogisten käsittemallien rakentamisessa ja vasta toisessa vaiheessa käsittemallien populoinnissa reaali maailman olioilla.

Omista ontologioista yhteiseen

Ontologioiden aihealueita ei kannata sekoittaa. On hyödyksi, että kullekin aihealueelle rakennetaan aluksi itsenäisesti sen reaali maailmaa kuvaava ontologia ja että maltetaan rajata pois rönsyt muille osa-alueille. Ei myöskään ole yhtä ja oikeaa ontologiaa aihealueelle, vaan riippuen näkökulmasta ja tarpeista ontologia voidaan rakentaa pätevästi myös uutena muunnelmana.

Eri aihealueiden ontologioiden kytkemiseen ja mahdolliseen yhdistämiseenkin löytyi yllättävän vaivattomat menettelyt. On XML-tekniikasta tuttua nimiavaruutta, Import-komentoa, SameAs-tägiä ja muitakin keinoja, joilla oman ontologiansa käsitteet linkittyvät muualla tehtyihin ontologioihin. Tuntuu mukavalta tietää, että Esim. suomalaisen yleisontologian FinnOnton kymmennettuhannet käsitteet suorastaan odottavat nurkan takana sitä hetkeä, kun haluan parilla klikkauksella linkittää jonkun niistä omaan käsittemalliini.

Kielen kiemuroita

Ontologian rakentaminen on tasapainoiluja filosofisen pohdiskelun ja käytännön ratkaisujen välillä. Jälkimmäisestä käy esimerkkinä nimeämiskäytäntö. Suomenkieli on toki helmi kielten joukossa, mutta omalla tavallaan hankala ontologian teossa. Pitkiä yhdyssanoja kannattaa välttää. Skandiaakoset näyttävät kelpaavan Protegelle ongelmitta, mutta XML-validointi käyttämälläni välineellä (XMLSpy) herjasi skandeista. En vielä tiedä, mikä käytännön merkitys tällä on, mutta pyrin itse nimeämään käsitteet ilman niitä. Skandeja voi riskittä käyttää Name- ja muissa lisämääritteissä, joten ei syytä huoleen.

Englanti on ontologiatyön "latina" kieliooppinsa ominaisuuksien ja monen ammattisanaston peruskielenä. Olen todella haasteellista mutta myös antoisaa pyrkiä työstämään kotikielistä ontologiaa. Olen nähnyt järkeväksi tehdä tavoitteessa kompromissin käsitteiden

relaatioiden kohdalla. Ilmaisen ne englanniksi, jolloin vältän suomenkielen sijamuotojen viidakon. Lopputuloksena on esim. väittämä "Lintu hasDestination Pesä", joka on ilmiselvästi sopivampi automatisoituun käsittelyyn kuin tyylikäs lause "Linnun tavoite on mennä pesäänsä".

Edellä kuvattua sekakielisyyden sekasortoa vähentää OWL-notaatioissa olennaisesti mahdollisuus liittää ontologian objektien, niiden ominaisuuksien ja keskinäisten suhteiden yhteyteen "annotaatioita" usealla kielellä. Esim. käsitteiden nimet annetaan "label-annotaatiolla" ja käsitteen selväkielinen selitys "comment-annotaatiolla" jokaisella halutulla kielellä. Protege ja xml-muunnosskeemat osaavat kaivaa esille ontologian halutun kielen versiona, mutta tämä vaatii jo erityistä taitoa, johon tämän kirjoittaja ei hevin tule pääsemäänkään.

Käsitteiden yksilöllisyys on järjestettävissä yllättävän luontevasti. Kullekin ontologialle määritetään yksiselitteiden URI selaindomainin avulla (esim. www.damafinland.org/concepts# ja www.damafinland.org/Sig#). Ainakin Suomessa Viestintävirasto kontrolloi niin tiukasti .fi tunnuksia, että edellä kuvaamallani domain-nimen käytöllä homma on hoidossa.

Työkaluista

Semanttisen webin tutkiminen on nyt suosittua yliopistomaailmassa. Tästä on ollut myös se myönteinen seuraus, että ontologioiden määrittelyjen ilmaistyyökaluja on tarjolla monia muitakin kuin kali-

fornialaisen Stanfordin yliopiston Protege. Se on aktiivisesti kehitetty open source-ohjelma, jonka ympärillä on melkoinen käyttäjäjyhteisö ja jolla on ei-kaupallisen java-ohjelmiston pieniä käytettyvyysongelmiakin.



Kuva 2. Herra Christian Wolffin itse.

Itse asiassa ahkera ja tarkka voi jopa Notepadilla kirjoittaa ontologiaa, onhan se pohjimmiltaan vain tekstimuotoista XML-skeemaan. Kaupallista tarjontaa on saatavilla niukasti. Niistä mainittakoon XML Spy-ohjelmasta tunnetun www.altova.com uunituore RDF/OWL editori SemanticWorks™ 2006, joka muutaman päivän käyttökokemuksella ei suppeassa ensi versiossaan vielä pärjää Protegelle.

Organisaation muilla tekniikoilla luotujen ja ylläpidettävien luetteloiden ja termistöjen liittäminen onnistuu OWL:lla ja Protegella näppärästi. Ne muunnetaan aluksi XML-muuntoskeemalla asiallisesti toisistaan riippumattomien "luokkien" dokumentiksi, joka täyttää muodollisesti ontologian muotovaatimukset. Sen jälkeen "luettelo-ontologia" lisätään vafrsinaisen ontologian nimiavaruuk-

siin ja tarvittaessa importeerataan fyysisestikin. Näin esim. yrityksen henkilöstöluettelo muuntuu www.firma.fi/persons# ontologiaksi, jolle annetaan nimiavaruustunnuksiksi nsjossa oleviin henkilöihin voi luoda viitauksen perusmuodossa "nsstaff=<http://www.firma.fi/persons>". Jo vain onnistuu onnistuu viitaus Heikkiin, tarvitaan vain ilmaisu "staff:Heikki".

Mukaan ontologioita luomaan

Suosittelen jokaiselle sukellusta kielen ja filosofian maailmaan. Varma tapa oppia tekemään ontologioita edullisesti ja mukavassa porukassa on tulla mukaan rakentamaan Dama Finlandin omaa ontologiaa. Toivottavasti mahdollisimman moni tarttuu tähän tilaisuuteen.

Pekka Valta työskentelee OKOn tietohallinnossa,
pekka.valta@oko.fi.

Osaamisyhteisöt

Systeemityöyhdistyksessä toimitaan niin yhdistystasolla kuin aihepiireittäin erikoistuneissa osaamisyhteisöissä. Monipuolisessa tarjonnassamme löytyy jokaiselle jotakin. Vaihtoehtona on myös perustaa omalle kiinnostukselleen uusi osaamisyhteisö - SYTYKE-hallitus toivottaa toimintaehdotukset tervetulleeksi. Osaamisyhteisön toimintaan pääset mukaan laittamalla postia vetäjälle.

Osaamisyhteisötarjontaan

kuuluu:

RELA keskittyy relaatiotietokantoihin, vetäjänä Lauri Pietarinen lauri.pietarinen@relationalconsulting.com.

ProjektiOSY— PrOSY pyrkii yhdistämään Systeemityön projektitoiminnasta ja sen kehittämisestä kiinnostuneet, vetäjänä Markku Niemi, markku.niemi@sttf.fi.

TestausOSY – FAST on testauksen keskustelu- ja yhteistyöverkosto, vetäjänä Maaret Pyhäjärvi, maaret.pyhajarvi@iki.fi.

DAMA Finland keskittyy tiedon, informaation ja tietämyksen hallintaan. Suomen osaston johtoryhmän (boardin) vetäjänä Pekka Valta, yhteyshenkilönä Minna Oksanen, minna.oksanen@gmail.com.

ViestintäOSY järjestää yhteistoimintaa viestintäsovellusten alueella, vetäjänä Tapani Ranta, tapani.ranta@generum.fi.

OlioOSY kehittää, tukee ja voimistaa oliomallilähtöistä sovel- luskehittämistä, vetäjänä Jukka Tamminen, jukka.tamminen@pp.inet.fi.

JavaSIG— on Javan käyttäjien ja harrastajien intressiryhmä, vetäjänä Simo Vuorinen, simo.vuorinen@tietoenator.com.

Tietämys ihmisen mentaalisenä ilmentymänä

Viljo K. Martikainen, TkT

Tarkastelen seuraavassa tietoa ja tietämystä lähinnä ihmisen tietoisien toiminnan kannalta. Samalla tavalla ontologista ongelmaa käsittelen pääosin ihmistieteiden ja filosofian näkökulmasta. Yhtymäkohtiakin voidaan löytää, mikäli käsitteiden rakenteet tietokoneiden ja palvelimien kovalevyillä saadaan sopivanasteiseen analogiseen suhteeseen ihmisen käsitteiden kanssa.

Termit eri yhteyksissä

ICT-alan globaali ongelma näyttää olevan mm. termien data, informaatio, tieto, tietämys referenttien ontologisen erottelun epämääräisyys. Tästä johtuneen niiden referentiaalisen ja myös niiden muiden merkitysten virheellinen ja usein täydellisen väärä tulkinta. Ongelman taustalta löytynee koulutusjärjestelmissämme ilmevä filosofian ensimmäisen kysymyksen ontologisen ongelman puutteellinen käsittely. Tällöin mm. kielen termien ja lauseiden referenttien ontologisen statuksen tärkeys jää vajaasti ymmärretyksi.

Sekaannusta ylläpitävät vankasti mm. monet suomenkielen termit,

kuten tietokirja, tietokone, tietolähde, tietomies jne. Kirjoissa on varmaankin kielen materiaalisina manifestaatioina informaatiota, jonka lukija tulkitsee omien käsitteittensä avulla elämyksellisesti koetuiksi kuvauksiksi ihmisistä, asioista, olioista, erilaisista toiminoista tai tapahtumista ja niiden merkityksistä ko. yksilölle hänen silloisessa elämäntilanteessaan.

Käsitteet ja mielen rakenteet

Väitöstyössäni olen tarkastellut ihmisen käsitteiden ja mielen systeemiä rakenteita ja niiden mahdollistamia uusia lähestymistapoja sekä ihmistieteiden että filosofian ikuisuusongelmiin. Käsitteet ovat olevan muistirep-

resentaatioita, joiden rakentuminen on aivojemme prosessuaalisten muistitoimintojen tuotosta. Meidän säilömuistimme tarjoaa mentaalisen tilan tai avaruuden koko elämäntilanteemme tärkeiden asioiden ja tapahtumien tallentamiseen. Näistä tallenteistahan muodostuu myös mieleemme muut rakenteet.

Työmuistimme puolestaan on se mentaalinen tila, jossa koemme kaikki tiedolliset, emotionaaliset, halu-, ja tahtoelementit, motivaatiomme ja myös kaikki erheiset elämyksemme, kuten harhat ja illuusiot. Sekä käsitteet että mieleemme rakenteet muotoutuvat jatkuvana dynaamisena prosessina

yksilöiden koko elämänkaaren aikana.

Ihmisen kokonaisuuden metamalli

Tiedon, tietämyksen ja informaation erottelun välttämättömyys ja hyödyllisyys selviää lukijalle vasta, kun tarkastelujen perustelu rakentuu niihin tieteellisiin faktoihin, joita mm. neuropsykologia, aivotutkimus ja empiirinen muistitutkimus ovat löytäneet ihmisen olemuksen, toiminnan ja maailmasuhteen selittämiseksi. Väitöstyöni muodostaa ihmisen kokonaisuuden metamallin, josta käytän nimitystä RA-malli. Se rakentuu ihmistieteiden tuloksina saavutettujen parhaiten selittävien teorioiden varaan.

Kriittisen tieteellisen realismin mukaan parhaiten selittävät teoriat lähestyvät kohteensa todellisuutta vastaavaa kuvausta tai ainakin toimivaa kuvausta. Tämä toimivuus on toisaalta pragmatistinen kannanotto tieteellisiin teorioihin.

Kuva1 (<http://lib.tkk.fi/Diss/2004/isbn9529929110/>) esittää RA-mallia eli ihmisen materiaalsen ja mentaalisen rakenteen ja hänen neuropsykologisen liittymänsä ydinkohdat. Niiden välityksellä ihminen on kytkeytynyt ympäristöönsä. Liittymän materiaalisesti manifestoituvina elementteinä ovat aistien koodaamat afferentit aktio-potentiaalit eli niiden sähkökemiallisten impulssien vuo, jonka aistielimemme koodaavat ulkoisista ja tai sisäisistä tapahtumista ja energialähteistä. Toisen impulssien vuon muodostavat aivomme motoriikkaamme ohjaavat osat,

jotka transformoivat päätöksemme ja valintamme puheentuottoa ja muuta toimintaamme ohjaavaksi efferenttien aktiopotentiaalien vuoksi.

Kokemuksen synty

Nuo kaksi problemaattista transformaatiota materiaalsen ja mentaalisen välillä ovat tieteelle vielä suuri ongelma. Vaikka eri aistien koodaaman informaation kulku aivoissa voidaan paikantaa tuhannesosasekunnin tarkkuudella, emme kuitenkaan tiedä, miten elämyksellinen kokemuksemme syntyy näistä miljoonien neuronien välisissä prosesseista.

Neuropsykologit ovat kuitenkin jo 1980-luvulla löytäneet visuaalsen informaation käsittelyn alueella ns. dorsaali- ja ventraalikanavat. Näistä edellinen voidaan nähdä online kanavana ja ventraalikanava muistikanavana, joka hakee aistinfon virittämänä tilannerelevantit käsitteet tulkitsemaan aistitun entiteetin eri merkityksiä. Tämän tulkinnan mukaan aivomme tulkitsee ulkoisen informaation ensin aistimuksiksi, joiden avulla olemme suorassa yhteydessä ulkomaailmaan tapahtumiin, joiden merkitykset kuitenkin tulkitsemme säilömuististamme noudettujen muistirepresentaatioiden eli käsitteiden avulla.

Informaation tulkinta

Käsitteet ovat tilannerelevantisti työmuistiin noudettuja kohteen kuvauksia, jotka selittävät ja täydentävät aistimuskuvausten sisältöä. Käsitteet ovat siten ne välineet, joiden avulla tulkitsemme kaikkea saamaamme informaatiota.

Käsitteiden avulla voimme nimetä kohteet, palauttaa mieleemme sen eri merkitykset ja jossain määrin ennustaa kohteen käyttäytymistä aikaisemman kokemuksemme perusteella.

Tieto on sellainen aivoprosessiemme muotoilema mentaalinen tila, jossa tilannerelevantien käsitteiden tulkitsemme ja ymmärrämme havaintokohteeseen liittyvät eri merkitykset. Näin ollen informaatio transformoituu tiedoksi vasta aivomme prosesseissa. Koska käsitteemme ovat yksilöiden ontogeneettisten historiain tuotoksia, niin infon tulkinta tiedoksi saattaa eri yksilöillä olla hyvinkin erilaista. Kuitenkin samoissa kulttuureissa ja työyhteisöissä eläneiden yksilöiden käsitteissä on varsin korkea intersubjektiivisuuden aste.

Tieto, tietoisuus, tietämys

Yllä olevan otsikon termien referentit voidaan ymmärtää paremmin, jos tiedämme millaisia ovat niitä selittävien käsitteiden olemus, rakenteet ja funktiot. Kuva 2 esittää väitöstyössäni kehittämäni ihmisen käsitteen systeemisen mallin. Sen sisärengas muodostuu kohde-entiteetin aistimellisista kuvauksista eli attribuuteista. Sisärenkaan keskellä on entiteetin nimi ja mahdolliset muut kielelliset kuvaukset, joilla viitataan siihen. Kaukoastien – kuulo, näkö ja haju – kuvaukset ovat renkaan oikealla laidalla, ja lähiaistien – maku, kosketus ja liike – kuvaukset ovat vasemmalla laidalla. Tilanteiden ja toimintojen kuvaukset ovat nimen yläpuolella ja muutosten ja reaktioiden kuvaukset nimen alapuolella.

Käsitteen ulkokehä muodostuu rakenteisista attribuuteista, jotka muodostetaan käyttämällä jo tunnettuja käsitteitä. Niistä muodostuu kolme saraketta, joista vasemmalla on käsitteen aikajana ja oikealla käsitteen teoriasarake. Keskelle jää käsitteen perus-, arvo- ja luokitusattribuuttien joukko, joihin kuuluvat myös kaikki sisärenkaan aistimelliset kuvaukset.

Rakenteisia attribuutteja entiteeteillä voi olla hyvin suuri määrä. Aistimellisten attribuuttien määrä on yleensä pienempi, eri aistilueidenkin kuvaukset voivat olla määrällisesti varsin suuria.

Käsitteen teoriasarakkeessa on entiteetin ontologisen statuksen, rakenteen ja funktioiden määrittelyt. Aikajanalla ovat entiteetin

historian, nykytilan ja ajatellun tai toivotun tulevaisuuden tilojen kuvaukset. Käsitteen luokitusattribuutit kuvaavat entiteetin niitä piirteitä, joiden perusteella se voidaan sisällyttää johonkin olioiden luokkaan. Entiteetin arvo- ja merkitysattribuutti kuvaa entiteetin aktuellia tai potentiaalista kykyä olla tai toimia subjektin jonkin toiminnan resurssina.

Time Line	Base and Value Line			Theory Line
FUTURE STATE	VALUE & MEANING			FUNCTIONS
The expected and possible states of the entity in the future according to subject's hopes or plans with probability calculations and timetables.	The value and meaning of an entity is defined by its ability to be or act as a resource or antiresource factor in subject's past, actual or future activities.			Are explanations of the purpose, innovations, abilities, exchanges, roles, and interactions of actors and the factual and potential uses and performances of non actors.
PRESENT STATE	MOTOR	SITUATIONS & ACTIONS	AUDITORY	STRUCTURE
	TACTILE	NAME or other linguistic expressions of the object entity	VISUAL	
	GUSTATORY	CHANCES & REACTIONS	OLFACTORY	
Situation description based on the actual sensory information and the recalled set of structural and other attributes illustrating the objects of attention.				The descriptions of the material and mental elements and their systems of the entity and their systems of the entity and its internal and external connections with other entities and their systems.
GENESIS	CLASSIFYING,			ONTOLOGICAL STATUS
The historical background, birth, and the development up to the present state of the object entity.	categorising, stereotyping, or dichotomic features of the object such as: good-bad, friend-enemy, opportunity – threat, eatable – uneatable, possible – impossible, coloured – white, succeeder – failure, ugly – beautiful etc.			The basic substance of the objects entity. Is it material or mental? Does it exist or subsist? What is its degree of intersubjectivity?

Kuva 2. The System Structure of Memory Representations and Concepts.

Käsitteen perusfunktio on toimia ihmisen ja maailman liittymänä, selittämällä koh-dattujen entiteettien eri merkitykset ko. tilanteessa. Käsitteiden monikäyttöisyys johtuu siitä, että jokainen aistimellinen attribuutti nimen ohella voi toimia ns. hakuattri-buuttina ja palauttaa yksilön työmuistiin kohde-entiteetin kaikki tilannerelevantit piirteet. Tavallisesti aivoprosessimme tuottaa tarvittavat mieleenpalautukset ilman suurempia ponnistuksia, mutta ei aina, kuten kaikki kokemuksesta tiedämme.

Ihmisen tieto ja tietoisuus selittyvät ihmisen RA-mallissa mentaalisina tiloina, jotka rakentuvat käsitteiden avulla ko. tilannetta vastaavalla tavalla käyttökelpoiseen muotoon. Ihmisyksilön tietämys on kuvattavissa kaikkien niiden tietojen ja taitojen koskevien muistirepresentaatioiden koosteena, jotka hän on ontogeneettisen historiansa aikana hankkinut ja ylläpitänyt mieleen ja käyttöön palautettavassa kunnossa. Yksilön tietämysvarasto on hänen toimintakykynsä mittari.

Apuja semanttiseen mallintamiseen?

Edellä kuvaamani ihmisen käsitteen rakenne avaa ihmistieteille ja filosofialle eräitä uusia tapoja tarkastella monia perinteisiä ongelmallisia kysymyksiä. Näitä ovat mm. tiedon, arvojen, kielen, merkien ym. teoreettisten entiteettien olemukseen liittyvät pulmalliset kysymykset. Voisiko käsitteen laajentaminen yksinkertaisesta määritelmästä moniulotteiseen rakenteeseen auttaa myös tietokoneisiin ja tietotekniikkaan liittyvien käsitteiden hyödyllisyyttä esim.

pyrittäessä rakentamaan semanttista verkkoa. Olisiko ohjelmoinnin olioiden ominaisuuksiin kytkettävä aistimellisten attribuuttien kyky toimia hakuattribuutteina? Miten tietokone ymmärtäisi ja ottaisi huomioon arvojen merkitykset sitä käyttävän ihmisen orientaatioiden mukaisesti? Onko lainkaan järkevää puhua koneen tietämyksestä, kun se ei kuitenkaan koe käsiteperustaisia elämyksiä? Taitaa se lopulta vain käsitellä ohjelmanmukaisesti siihen syötettyä informaatiota ymmärtämättä juuri paljokaan sen merkityksistä. Nehän syntyvät vasta ihmisen arvioinnin jälkeen hänen mielessään, kun hän punnitsee tietokoneen välittämän informaation merkityksiä oman tilanteensa ja arvojensa mukaisesti.

Ontologian käyttöön liittyviä ongelmia

Ontologiajakäsitetermienerilainen käyttö ihmistieteissä ja ICT-alalla voi myös olla ongelmia synnyttävää ei alan kieleen vihkiytyneelle. Tietokoneen käsitteissä ei ole välittömästi koettuja aistimellisia attribuutteja ellei konetta kuljetavan robotin näkö-, kuulo- haju-, tunto-, maku-, ja liike-'aistielinten' prosessoimia tuotoksia saada toimimaan hakuattribuutteina niiden toimintaa ohjaavien käsitteiden eli teoreettisten entiteettien attribuuteiksi, joihin kytketyt linkit ja valintasäännöt ovat mielekkäitä ko. tilanteessa. Tilanneorientoituvasti ja mielekkäästi toimiva tietokone ja sen ohjaama robotti lienee vielä kaukana tulevaisuudessa, vaikka alkeelliset sellaiset jo imuroivatkin lattioita.

Kuva 1 on väitökirjassani sivulla 30 ja kuva 2 sivulla 69. Väikkäri on verkossa TKK v. 2004 väitöskirjat. Linkki on: <http://lib.tkk.fi/Diss/2004/isbn9529929110/>

Olen valmistunut TKK :sta 1958. Olen nykyään eläköitynyt liikkeenjohton konsultti. Opetan kuitenkin johtamisen filosofiaa TKK:lla avoimen yliopiston puolella. Väittelin vasta kypsässä iässä (TKK 2004) aiheista, joista olen "saarnannut" 40 vuotta johtoryhmien seminaareissa erilaisten organisaatioiden kehittämisohjelmien ja erilaisten konferenssien yhteydessä. Väitöskirjani Concepts and Mind as Dynamic Memory Systems Structuring the Human Mental on luettavissa Internetissä. Se on monitieteinen, kokonaisvaltainen esitys ihmisen mentaalisen maailman olemuksesta, rakenteista ja funktioista, jotka kaikki rakentuvat empiirisen aivo- ja muistitutkimusten eräisiin keskeisiin tuloksiin.



Hiljainen tieto yrityksen voimavarana

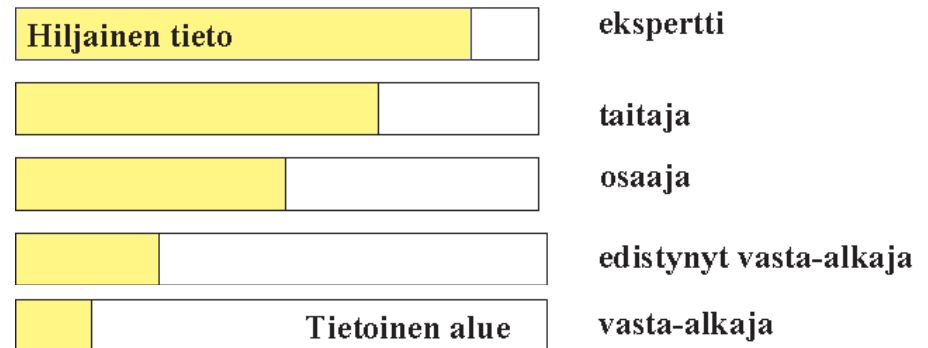
Paula Miinalainen, Oy Arbor Vitae Finland Ltd

Pettynyt huokaisu ”liikaa yrittämistä” on varsin tuttu ilmaisu golfkentällä. Hyväkin pelaaja saattaa joskus pilata lyöntinsä ”yrittämisellä”. Yrittäessään hän kohdistaa huomionsa lyöntiin tietoisesti ja aiemmin automaattisesti, itsestään sujunut hyvä lyönti rikkoutuu ja pallo menee ties minne. Itsestään sujuva hyvä lyönti tapahtuu hiljaisen tiedon varassa, tietoinen lyönnin tarkkailu muuttaa osaamistasoa alhaisemmaksi.

We know more than we can tell

Unkarilainen lääketieteen ja kemian tutkija Michael Polanyi huomasi 1940-luvulla, että meillä on sellaista tietoa, jota emme osaa pukea sanoiksi. Se on kokonaisvaltaista, kokemukseen perustuvaa tietoa. Siihen sisältyy kaikki se geneettinen, ruumiillinen, intuitiivinen ja kokemusperäinen tieto, jota meillä on. Kaikki se informaatio, joka aivoissamme on, ja josta vain hyvin pienestä osasta olemme tietoisia.

Polanyi määritteli tiedon ja taidon kolmeen tasoon. Ensimmäinen taso on tietoa, josta osaamme puhua. Sitä voidaan kuvata kielen avulla. Toinen taso on se artikuloimaton taso, jota osataan kuvata kielellisesti vain kun siihen kohdennetaan erityistä huomiota. Muutoin



Kuva 1. Osaamisen tasot. Tutkijat Stuart Dreyfus ja Hubert Dreyfus.

se on hiljaista tietoa, joka tapahtuu itsestään. Kolmas taso on hiljainen tieto (tacit knowledge). Se on se osa tiedosta, jota ei osata kuvata lainkaan kielellisesti.

Osaamisen tasot

Tutkijoiden Stuart Dreyfus ja Hubert Dreyfus (1999) mukaan kaikessa osaamisessa on mukana myös hiljaista tietoa. He jakoivat

osaamisen viiteen tasoon. Ne ovat vasta-alkaja, edistynyt vasta-alkaja, osaaja, taitaja ja ekspertti.

Vasta-alkaja osaa sijoittaa oppimansa käsitteet käytännön tilanteisiin.

Hän käyttää step-by-step sääntöjä. Vasta-alkajalla voi olla paljon teoreettista tietoa, mutta ei vielä

tietoa siitä, miten tätä tietoa tulee soveltaa. Hiljaisen tiedon osuus on vähäinen.

Edistynyt vasta-alkaja päätelee kussakin tilanteessa, miten kannattaa toimia. Hän käyttää step-by-step sääntöjen lisäksi tilannekohtaisia sääntöjä. Hän keskittyy edelleen yksityiskohtiin. Hän tarvitsee vielä tukea ja haluaa ohjeet ja säännöt kirjallisina. Hän ei vielä hahmota kokonaiskuvaa, ei vielä ymmärrä kaiken suhteellisuutta. Hiljaista tietoa on kuitenkin jo enemmän kuin vasta-alkajalla.

Osaaja ymmärtää että kussakin tilanteessa vaikuttavia tekijöitä on paljon. Osaaja tekee suunnitelman tai valitsee tietyn näkökannan tai päämäärän hallitakseen suurta vaikuttavien tekijöiden määrää. Hän voi laatia suunnitelmaan sopivia uusia sääntöjä. Hän huomaa kuitenkin, että sääntöjä ei voi laatia kaikkeen, mikä eteen tulee. Kaikkea ei voi tietää etukäteen ja hän joutuu tekemään valintoja ja päätöksiä riittämättömin tiedoin ja kuitenkin toimimaan. Osaaja näkee toimintansa osana pitkäaikaista suunnitelmaansa. Peilaamalla asioita suunnitelmaan hän päättää mikä on tärkeää mikä ei. Osaajalla tämä on tietoisista pohdintaa. Osaaja on joustava ja tunnistaa toiminnallisia malleja. Osa osaamisesta on hiljaista tietoa.

Taitaja tarkastelee tilannetta kokonaisuutena. Näkökulma tilanteeseen ei perustu tietoiseen harkintaan vaan se ”avautuu” aikaisempien kokemusten kautta. Suunnitelmat syntyvät intuitiivisesti, ikään kuin itsestään. On helpompaa nähdä,

mikä tilanne on kuin tietää, mitä pitää tehdä. Taitajalla on toimintavaihtoehtoja paljon. Hän osaa laittaa osat tärkeysjärjestykseen. Laajasta kokemuksesta kumpuaa ideoita, mielikuvia ja tunteja, joita kaikkia ei osaa pukea sanoiksi. Hiljaisen tiedon osuus on lisääntynyt.

Ekspertti näkee kokonaistilanteen ja toimii syvällisen ymmärtämisen pohjalta. Kysymykseen miksi teit noin ekspertti vastaa: Se vain tuntui siltä tai Se vaikutti hyvältä tai En osaa eritellä tai Se nyt näytti toimivalta. Intuitiivinen kokemus ohjaa toimintaa ja tietoiset säännöt ovat mukana vain osittain. Ekspertti osaa mennä suoraan ytimeen. Hän huomio olennaisen. Hän luottaa intuitioon. Toiminta on ajatuksetonta hallintaa niin kuin kävely. Hiljaisen tiedon osuus on suuri.

Hiljainen tieto yrityksessä

Kirjassa Strateginen kumppanuus (Kari Laento ja Pirjo Stähle) todetaan, että yrityksen tietopääomasta 95 % on hiljaista tietoa, vain 5 % on dokumentoitu. Yrityksissä arvostetaan vain tätä dokumentoitua tietoa. Tämä 95 % on mielestäni yllättävän suuri osuus, koska prosessien kehittämiseen ja niiden kuvaamiseen on yrityksissä panostettu paljon. Kuitenkin saattaa käydä niin, että hiljaisen tiedon merkitys huomataan vasta kun se on menetetty henkilön siirryttyä muualle. On tärkeätä tunnistaa yrityksen hiljainen tieto ja tunnistaa myös huippuosaajat. Kun hiljainen tieto saatetaan tietoiseksi, niin siitä syntyy arvostusta ja tieto nähdään aidosti voimavarana. Myös epäonnistumiset ovat merkittäviä tietolähteitä. Niistä saatu kokemus

kannattaa dokumentoida hyötykäyttöön.

Hiljainen tieto ja muutos

Yritysten tulisi ensin tulla tietoisiksi kulttuurissaan olevasta hiljaisesta tiedosta ennen kuin kulttuurissa voi tapahtua muutoksia. Englantilainen talouselämän tutkija David Gurteen sanoo, että muutos alkaa yksilöstä. Yksilön täytyy tulla tietoiseksi omasta hiljaisesta tiedostaan voidakseen hyväksyä muutoksen ja muuttua. Kun yksilö tulee tietoiseksi osaamisestaan, hän voi sitä itsekin arvioida ja pohtia tietoisesti tarvittavia muutoksia toiminnassaan. Tätä kautta muutos voi onnistua.

Uusien järjestelmien ja teknikkoiden käyttöönoton suunnittelussa tämä on olennainen tieto. Hiljaisen tiedon tiedostaminen on välttämättömän vaihe. Tätä kautta voidaan ymmärtää muutosvastarinta uudesta näkökulmasta. Ne henkilöt, mahdolliset huippuosaajat, joilla on paljon hiljaista tietoa, tarvitsevat aikaa olemassa olevan hiljaisen tiedon tiedostamiseen. He voivat oppia uuden toimintatavan vasta tämän tiedostamisvaiheen jälkeen. Vasta-alkajilla ei tätä tarvetta ole.

Pari esimerkkiä omista kouluttajakokemuksistani. Koulutin 90-luvun alkupuolella tekstinkäsittelyohjelman WP:n (Word Perfect) käyttäjille Windowsin Wordiä. Koulutin useita satoja henkilöitä ja havaitsin, että alussa Word oli kaikkein hankalinta sellaisille oppijoille, jotka osasivat erinomaisesti WP:tä. WP:llä heidän kädet toimivat nopeammin kuin ajatus. Se oli sitä hiljaista tietoa. Heidän

täytyi ensin oppia ajattelevaan tietoisesti kirjoittamista, jotta uudet toiminnot saattoi ottaa käyttöön. Vastaavanlainen kokemus minulla on ohjelmointikielen opettamisesta.

Oppijaryhmässä olivanhajaeksperttejä ja juuri korkeakoulusta valmistuneita. Ekspertit, vanhat konkarit olivat suunnitelleet ja ohjelmoineet vuosia järjestelmiä konekielellä. Nyt piti oppia aivan toisenlainen kieli. Korkeakoulusta tulleet näyttivät omaksuvan opittavan kielen nopeasti ja tekivät harjoitukset vaivatta. Konkarit tuskailivat.

Minulla oli mahdollisuus seurata tätä oppimisprosessia pitempäänkin. Sillä koko oppijaryhmä jatkoi suunnittelu- ja ohjelmointityötä projektissa, jonka projektipäällikkönä toimin. Konkareiden tuska jatkui vielä jonkin aikaa, sillä he aivan ilmeisesti kirjoittivat ohjelmaa päässään vielä vanhan kielen pohjalta ja joutuivat umpikujaan. Kielten rakenteethan olivat aivan erilaiset. Heidän siis tuli ensin tulla tietoiseksi tästä vanhasta tavasta, siitä hiljaisesta tiedosta ja sen jälkeen rakentaa uuden ohjelmointikielen mukaiset rakenteet. Mutta kolmen kuukauden kuluttua konkarit olivat jo saavuttaneet erinomaisen osaamisen tason myös tässä uudessa ohjelmointikielessä ja samalla heillä oli käytössään kaikki muu kertynyt ammattitaito. He olivat jälleen eksperttejä työssään.

Hiljainen tieto ja yrityskulttuuri

Yrityskulttuuri kehittyy pitkän ajan kuluessa. Kulttuuriin kuuluu yrityksen tavat ja tottumukset. Siihen kuuluvat myös yrityksen

arvot. Kulttuuri määrittelee miten työpaikalla tulee toimia, mikä on hyväksyttyä ja mikä ei ole. Uudet työntekijät kasvavat kulttuuriin sisään. He oppivat perehdyttämisen kautta ja ennen kaikkea työkavereilta työtä tehdessä. Näin hiljainen tieto siirtyy hiljaisesti. Joskus kulttuuri voi olla niin voimakas, että muutokset ajatellaan mahdottomiksi. ”Onko tämä kiveen hakattua?” on itse asiassa hyvä kysymys.

Ruotsissa on hiljaista tietoa tutkittu Arbetslivsinstitutissa kaksikymmentä vuotta. Tämän tutkimuksen mukaan hiljaista tietoa ei sinänsä voida siirtää, mutta siitä voidaan tulla tietoiseksi. Kun hiljaisesta tiedosta tullaan tietoiseksi niin voidaan arvioida sen hyödyllisyyttä. Hiljainen tieto voi olla myös vanhentunutta, nykyiseen tilanteeseen sopimatonta. Se voi olla joissakin tapauksissa luonteeltaan syrjivää. Hiljainen tieto voi olla myös huippuosaamista. Se voi olla oikea yrityksen salainen ase. Huipputaitajien toimintatapaa voidaan mallintaa, jolloin tavoitteena on mallintajan avulla purkaa huippuosaajan taidot auki. Mallinnuksen pohjalta luodaan toimintamalli, joka voidaan opettaa muille.

Pohdimme Helsingin tietojenkäsittely-yhdistyksen Akva-seminaarissa ryhmätyönä hiljaisen tiedon tunnistamista ja sitä, miten se saadaan esiin. Tässä joitakin vinkkejä ryhmätöiden pohjalta. Hiljaista tietoa on henkilöllä, johon viitataan usein sanomalla ”Kysy siltä.” Työyhteisöissä tällaisia henkilöitä on aina. Hiljainen tieto voidaan

löytää myös kyselemällä, kuuntelemalla, kehityskeskusteluissa, asiaspalauteista jne.

Miten hiljaista tietoa saadaan yrityksissä esiin? Yrityksen kannattaa luoda oma hiljaisen tiedon strategia. Kun hiljaista tietoa arvostetaan, niin syntyy myönteinen ilmapiiri. Huomataan, että asiantuntemus on meillä itsellä eikä ulkopuolisilla. Hiljaista tietoa voidaan synnyttää ja levittää yhteisissä keskustelutilaisuuksissa, luomalla sparrausporukoita, joissa on yhdessä eri tasoisia osaajia konkareista vasta-alkajiin. Voidaan luoda hiljaisen tiedon mittarit. Ryhmätöissä tuli esiin aivan loistava idea. Valitaan ja palkitaan VUODEN TYÖTOVERI. Tarkoituksena on oman osaamisen tunnistaminen ja tiedostaminen. Sitten sen työkaverille jakaminen sekä tiedonkulun parantaminen.

Näistä kilvoitellaan.

Lähteet

Veli-Matti Toivonen, Riitta Asikainen: Yrityksen hiljainen osaaminen

Kari Laento ja Pirjo Stähle: Strateginen kumppanuus

Hetkyn Akvaseminaarin
15.10.2005 ryhmätöyt

Toimitusjohtaja Paula Miinalainen
Oy Arbor Vitae Finland Ltd:stä, www.arborvitae.fi, on toiminut kaksikymmentä vuotta sekä IT-konsulttina, projektipäällikkönä ja suunnittelijana että projektinohjauksen, systeemyön, testauksen ja laadunvarmistuksen kouluttajana.

paula.miinalainen@arborvitae.fi
puhelin 050 500 2363.

Arkkitehtuurit - taidetta ja tekniikkaa

Tarja Raussi, Sytyke ry

Järjestyksessään 8. Sytyke-seminaari kokosi innokkaan ja keskusteleavan joukon jäsenistöä Viking Mariellalle 7.-9.9.2005. Tänä vuonna teemana oli arkkitehtuurit, joista kuulimmekin monipuolisia esityksiä.

UML:ää pitää osata soveltaa!

Puheenjohtaja Helena Venäläisen avauksen jälkeen Jouko Poutanen IBM:ltä pääsi kertomaan, miten saadaan ohjelmistokehitykseen tehoja arkkitehtuurimallinnuksesta. Ohjelmistokehityksen nykyisiä haasteita ovat mm. välineiden runsaus, ohjelmistokehitystapojen muuttuminen sekä liikkuvien osien määrän kasvu. Mallinnus on avain tämän monimutkaisuuden hallintaan. Mallikeskeinen kehitys (Model Driven Development) tarkoittaa liiketoimintalogiikan ja tuotannon käytäntöjen kuvaamista malleilla, joita käytetään toteutuksessa, testauksessa ja ylläpidossa. Jouko painotti sitä, että UML:ää pitää osata soveltaa: mitä elementtejä tarvitaan ja miten niitä projektissa sovelletaan. Yhdistellään

kaaviotyyppejä, jotta ilmaisuvoima kasvaisi, esimerkiksi samaan kaavioon laitetaan käyttäjärooleja, solmuja, yhteistyö- eli collaboration-symboleja. Abstraktiotason nostaminen alentaa kustannuksia ja nostaa tehokkuutta. Jouko havainnollisti IBM:n Rational Software Architect -tuotteella, miten MDD-kehittämistä voidaan tehdä.

Arkkitehtuurikehyksen

näkökulmat apuna

monimutkaisuuden hallinnassa

Pekka Kähkipuro Sysopen Digiasta kävi läpi mobiilien yritysjärjestelmien haasteita ja ratkaisuja liiketoiminnan, arkkitehtuurien ja osaamisen näkökulmasta. Mobiili-integraatio tarkoittaa sellaisten ICT-ratkaisujen kehittämistä,

jotka ulottuvat mobiiliedustasta (front-end) taustajärjestelmiin (back-end). Ideana on siirtyä erilisistä sovelluksista prosessien välisiin tapahtumiin. Jotta tämä onnistuisi, jokaisella liiketoimintaprosessilla on oltava selkeä omistaja. Pekka esitteli arkkitehtuurikehyksen (Architecture Framework) eri näkökulmia ja niiden avainkäsymyksiä, joihin pitää etsiä vastaukset. Paljon pohjatyötä ja perusasioiden selvitystä sekä päätöksiä on tehtävä, jotta mobiili-integraatio onnistuisi. Arkkitehtuurikehyksen näkökulmien käyttö auttaa monimutkaisuuden hallinnassa.

Innovaatiot ja standardointi tasapainoon

Eero Koskinen SAP Finlandilta kertoi integraatioarkkitehtuurin toteutuksesta SAP NetWeaverin avulla. Haasteena on innovaation ja standardisoinnin tasapainottaminen ts. miten tuottaa halvalla ja tehokkaasti ja samalla mahdollistaa innovatiiviset ratkaisut. Ideana on, että yrityksen järjestelmät ja palvelut voidaan kattaa ns. yhdistelyalustalla (composition platform), jollainen on esim. SAP NetWeaver. Tämän päälle voidaan sitten rakentaa yhdistelysovelluksia (composite applications), jotka ohjaavat prosesseja.

Kokonaisuus on yhtä vankka kuin heikoin lenkki

Tomi Tuominen TietoEnatorista kertoi www-sovellusturvallisuudesta hyvin havainnollisesti sekä antoi vinkkejä, miten asiat pitää

hoitaa. Tärkein niistä oli, että syöte ja tulos pitää aina tarkistaa. Muista mainittakoon esim. "sipu-
limallin" hyödyntäminen (defense in depth) ja asioiden pitäminen yksinkertaisina. Peruseriaatteiden vahvistukseksi hän esitti ohjelmakoodipätkien avulla, kuinka voi käydä, jos asioita ei tehdä oikein. Www-tietoturvaan liittyvää aineistoa löytyy esim. osoitteesta www.owasp.org.

Liiketoimintaprosessien ymmärtäminen avainasia **SOAssa**

Sakari Olli Tieturista pohti, onko SOA ajattelutapa vai tekniikkaa. Vai onko kyse vanhoista tutuista asioista, joille on annettu uusi trendikäs nimi? Tavoitteena on rakentaa entistä selkeämpiä, korkean tason palveluita liiketoiminnalle tarjoavia järjestelmiä. Liiketoimintaprosessien ja niiden vaatimusten ymmärtäminen on tässä avainasemassa. Korkeamman tason palvelujen luominen

tarkoittaa käytännössä eri järjestelmien integrointia toisiinsa. Eri tekniikoilla tehtyjä sovelluksia on voitu yhdistää ennenkin esim. CORBAlla, mutta se on ollut kallista. Nykyään on helpoja keinoja, kuten WebServices, joilla liimata sovelluksia toisiinsa. Integraatiomaailmassa on tärkeää abstraktiotason nostaminen: käyttäjä ei tiedä, mistä järjestelmästä tai niiden yhdistelmästä hänen saamansa palvelu on peräisin. Puheenvuoro herätti runsasta keskustelua mm. XML:n raskaudesta mobiilipalvelujen toteutuksessa.

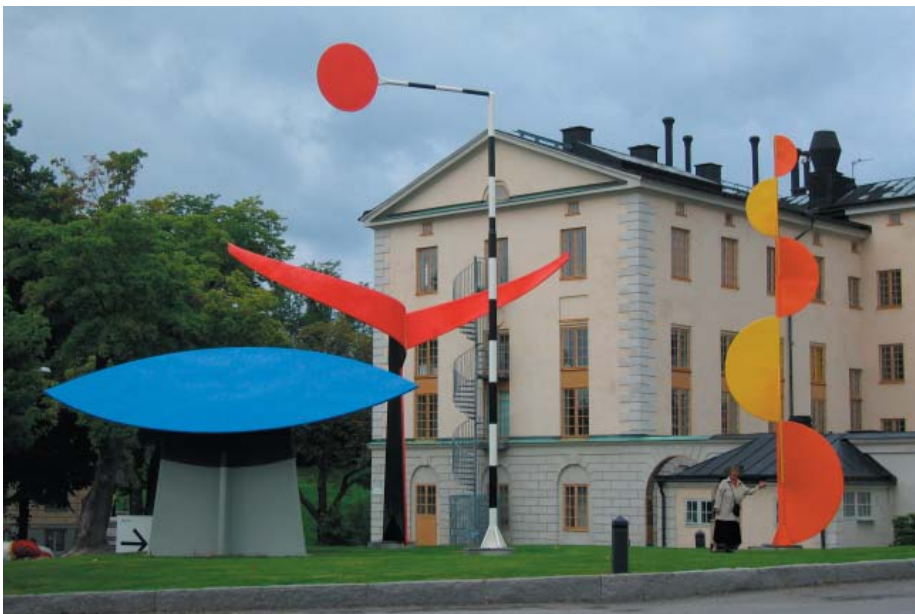
Ensimmäisen päivän päätteeksi seminaarin sponsorit saivat kahden minuutin puheenvuoron, jonka jälkeen osallistujat vaihtoivat vapaalle. Kollegoihin tutustumista ja verkostoitumista jatkettiin buffet-illallisella. Torstai-aamuna saavuimme Tukholmaan, jossa osallistujilla oli vapaata oleskelua klo 15 asti. Tukholmassa näimme

monenlaista arkkitehtuuria, niin kammutuksia kuin hienouksiakin.

Teknologia on vain mahdollistaja

Toisen päivän avasi Jari Isokallio TietoEnatorista aiheella Yritysarkkitehtuurit: hypeä vai asiaa. Yritysarkkitehtuurin (EA) tarkoituksena on poistaa IT:n ja liiketoiminnan välinen kuilu. EA sisältää liiketoiminta-, informaatio-, järjestelmä- ja teknologia-arkkitehtuurit. Ei riitä, että EA vain kehitetään, se pitää ottaa käyttöön sekä sitä pitää ylläpitää ja hallita. EA:n pitää olla tarpeeksi hyvä ja hyödyllinen, mutta ei täydellinen. Teknologia on vain mahdollistaja, ei itseisarvo. Puheenvuoro herätti runsasta keskustelua, mm. pitäisikö tietohallinnon vai liiketoiminnan vetää EA:n kehittämistä?

Aali-Petri Alikoski Microsoftilta puhui aiheesta Software Factories: järjestelmien mallinnus Microsoftin välineillä. Sovellusarkkitehtuurin ja infra-arkkitehtuurin välillä voi olla viestintäongelmia, mm. vastuiden jako on epäselvä. Infran ylläpidon tuntemus vähentää väärinymmärryksiä. Työkalujen pitää vähentää monimutkaisuutta, olla toisiinsa integroituvia sekä sopia sekä käyttäjäorganisaatiolle että riippumattomalle ohjelmistotoimittajalle. Visual Studio 2005 sisältää useita erilaisia versioita eri käyttäjärhyille. Software Factories tarkoittaa tapaa suunnitella ja toteuttaa DSL-pohjaisesti järjestelmiä. (DSL = Domain Specific Language) Aali kertoi, että Microsoft on kehittänyt aivan oman DSL-kielensä. Visual Studio ei tule sisältämään kaikkia UML-kaavioita.



Kuva 1. Moderna museet oli yhdistänyt tekniikan ja taiteen, sillä taideteokset liikkuvat! Rakennus itse edusti perinteisempää arkkitehtuuria. Etsi kirjoittaja kuvasta. Kuva: Minna Oksanen.

Arkkitehti ei saa olla poliisi

Aija Palomäki Nokiasta kokosi hyvin yhteen seminaarissa esitellyjä arkkitehtuureja omassa puheenvuorossaan informaatio-arkkitehtuureista. Informaatioarkkitehtuurissa on kyse koko yrityksen tietomallista: sen peruskäsitteistä, niiden sisällöstä ja välisistä suhteista. Hankekohtaisten tietomallien pitää olla linjassa yrityksen malliin. Esimerkkinä hän käytti Nokian omaa informaatioarkkitehtuuria, joka sisältää geneerisiä malleja, joissa ei ole liiketoimintalueiden yksityiskohtia. Hän myös kertoi kokemuksia informaatioarkkitehtuurin käyttöönotosta ja

siitä saadusta palautteesta. Aija korosti, että yritystason arkkitehtuurien kehittäminen pitää aina perustella hyvin, sillä niiden kehittäminen maksaa. Perusmäärittelyt pitää hyväksyttää laajalla joukolla, muuten ne eivät tule koskaan käyttöön.

”Jo aate EAI toimi”

Helena Venäläinen totesi seminaarin yhteenvetokeskustelussa, että ainakin päivien aikana on käynyt selväksi, miten monipuolisesta ja -tahoisesta asiasta arkkitehtuurissa on kyse. Päivän lopuksi osallistujat nauttivat a la carte -illan, jonka aikana pöytäseurueet

kilpailivat keskenään seminaariin liittyvästä iskulauseesta, joka piti muodostaa pöytäseurueen nimien alkukirjaimista.

Tarja Raussi on systeemityön asiantuntija Tieturi Oy:ssä, erikoisalueenaan mm. vaatimusmäärittely ja UML:n käyttö eri systeemityön tehtäväkokonaisuuksissa. Hän on myös Sytyke ry:n johtokunnan jäsen.

Hyvä Tietotekniikan liiton jäsen!

Päivitä tietosi

Voit päivittää jäsentietosi verkkosivuillamme www.tt-tori.fi. Tietojen päivittämiseen tarvitsen käyttäjätunnuksen (= jäsennumerosi, merkitty jäsenlehtiin) ja salasanasasi (=postinumerosi). Jos olet muuttanut salasanasiasi tai kirjautuminen ei muutoin onnistu, voit lähettää tunnusten tarkistuspyynnön osoitteella jasenasiat@ttlry.fi.

Toivomme sinun erityisesti varmistavan, että sähköpostiosoitteesi jäsentiedoissa on oikea.

Henkilökohtaisempaa palvelua – Sinun eduksesi

Tietotekniikan liitto jäsenyhdistyksineen, osaamisyhteisöineen ja kerhoineen haluaa palvella jäseniään henkilökohtaisemmin ja paremmin, tarjota tietoa juuri Sinua kiinnostavista aiheista. Palvelun parantamiseksi uusimme verkkopalvelumme helmikuun 2006 lopussa. Uuden palvelun avauduttua voit auttaa meitä räätälöimään palvelumme juuri Sinulle. Päivität vain tiedot itseäsi kiinnostavista aiheista ja saat tietoa juuri niistä. Voit päivittää valintasi aina halutessasi. Tietoja ei anneta ulkopuolisille tahoille vaan niitä käytetään ainoastaan TTL:n ja sen piirissä toimivien yhteisöjen tarkoituksiin.

Tietotekniikan liitto ry
Lars Sonckin kaari 12
02600 Espoo
p. 020 741 9898
f. 020 741 9889
www.tt-tori.fi
etunimi.sukunimi@ttlry.fi
jasenasiat@ttlry.fi

SYTYKE ry on vuodesta 1987 toiminut valtakunnallinen systeemityöntekijöiden ammatillinen yhdistys, joka kehittää alan ammattilaisten välistä yhteistyötä ja tutkimustoimintaa.

Teemayhdistyksen jäseneksi voivat liittyä kaikki systeemityöstä kiinnostuneet yksityiset henkilöt, yhdistykset ja yritykset. SYTYKE ry:n toiminta-alueena on koko Suomi. SYTYKE on Tietotekniikan liitto Ry:n jäsenyhdistys.

Lisätietoja SYTYKE ry:stä: <http://www.sytyke.org/>

Johtokunta 2006

SYTYKE ry:n johtokunnan sähköpostilista: sytyke-hallitus@pcuf.fi

www-sivut: <http://www.pcuf.fi/sytyke>

Helena Venäläinen (puheenjohtaja)

Osuuspankkikeskus
PL 308
00101 Helsinki,
Puh. +358 010 252 3690;
+358 50 568 6690
Fax +358 010 252 3170
helena.venalainen@op.fi

Kari Uusi-Äijö

Pohjolan Systeemipalvelu Oy
Puutarhurinkuja 2
00300 Helsinki
Puh: +358 40 835 6541
Fax: +358 10 559 3906
kari.uusi-aijo@pohjola.fi

Ilkka Pirttimaa

Stockmann
Kutomotie 1C
PL 147
00381 HELSINKI
Puh. +358 50 389 0022
ilkka.pirttimaa@stockmann.fi

Markku Niemi

STTF Oy (Software Technology Transfer Finland Oy)
Tekniikantie 14, 02150 Espoo
Puh. +358 50 5124 687
Fax 09 - 80 45 13 33
markku.niemi@sttf.fi

Tarja Raussi

Tieturi Oy
HTC Santa Maria,
FIN-00180 Helsinki
Puh. +358 (0)9 431 551
Fax +358 (0)9 4315 5302
tarja.raussi@tieturi.fi

Jori Rätty

SysOpen Digia Oy
Hiomotie 19, FIN 00380
Helsinki
Puh: +358 50 551 5152
jori.ratty@sysopendigia.com

Paula Miinalainen

Arbor Vitae
Mechelininkatu 16 A 9
00100 Helsinki
Puh. +358 50 500 2363,
paula.
miinalainen@arborvitae.fi

Lauri Laitinen (varajäsen)

lauri.laitinen@nokia.com

Lea Virtanen (varajäsen)

lea.virtanen@kolumbus.fi

Liittokokous- edustajat

Silja Räisänen

silja.raisanen@pohjola.fi

Jori Rätty

jori.ratty@sysopendigia.com

Helena Venäläinen

helena.venalainen@op.fi

Kati Viik

kati.viik@helsinki.fi

Toimisto

Puhelinvastaus- ja sihteeripalvelu VT Oy/Susanna Koskinen
Systeemityöyhdistys Sytyke ry
Henrikintie 7 A
00370 Helsinki
Puh. 09-5607 5363
Fax 09-5607 5365
sytyke@hennax.fi

Sytyttääkö? - Liity jäseneksi

Systeemityöyhdistyksen jäseneksi liitytään Tietotekniikan liiton kautta (<http://www.tt-tori.fi/>, (09) 4765 8530, jasenasiat@ttlry.fi) valitsemalla jäsenyhdistykseksi Systeemityöyhdistys ry. Nykyinen Tietotekniikan liiton jäsen voi liittyä joko vaihtamalla jäsenyhdistystä tai liittymällä lisäjäseneksi.

Tietotekniikan liiton henkilöjäsenmaksu vuonna 2006 on alkaen 52 €, erityisryhmien hinnoittelusta lisätietoja Tietotekniikan liitosta. Lisäjäsennyys maksaa 11 €/yhdistys.



Toteutuuko liiketoiminta IT-järjestelmissäsi?

Liiketoiminnan ja IT:n välinen yhteys realisoituu organisaation prosesseissa. Prosessit toimivat järjestelmäkehityksen peruspilarina. Liiketoimintälähtöisyys ei kuitenkaan automaattisesti siirry järjestelmiin. Se edellyttää tietoista paneutumista kehittäjiltä liiketoimintaan mm. vaatimusten ja arkkitehtuurin määrittelyssä. Toisaalta teknologia tuo uusia mahdollisuuksia myös liiketoimintaan. Ota selvää miten.

Tieturin asiantuntijat sekä laaja ja monipuolinen tarjonta on käytössäsi.

Ota yhteyttä ja kysy lisää palveluistamme

Myyntijohtaja Jukka Tuominen
(09) 4315 5260, jukka.tuominen@tieturi.fi

Johtaja Päivi Hietanen
(09) 4315 5664, paivi.hietanen@tieturi.fi

www.tieturi.fi/ohjelmistotuotanto

Ohjelmistotuotanto

- Prosessien mallintaminen 23.1.Tampere
- Moderni systeemyö 6.2.
- Järjestelmävaatimusten hallinta 15.-17.2.
- Oliot ja UML määrittelijälle 19.-20.1.
- Arkkitehtuurin suunnittelu (UML) 23.-24.1.
- Oliomäärittely ja -suunnittelu (UML 2.0) 1.-3.2.
- XML-kieli 30.-31.1.
- Web Services- ja SOA-yleiskatsaus 20.2.
- Java2-järjestelmän projektipäällikkövalmennus 1.-3.2.
- .NET-järjestelmän projektipäällikkövalmennus 13.-15.2.
- Järjestelmäintegroinnin valmennusohjelma 27.2. alkaen
- Testauksen valmennusohjelma 23.-25.1.
- ISEB Foundations Certificate -valmennusohjelma 13.-15.2.
- ISPL Foundations 15.-17.3.

Ajankohtaistapahtumat

- ITIL Foundations 24.-26.1. ja 13.-15.2.
- ERP 2006 1.2.
- Open Source 2006 1.2.
- Testaus 2006 1.2.

Ilmoituksessa olevien kurssien sisällöt ja hinnat löydät kätevimmin osoitteesta www.tieturi.fi/sytyke.

Ilmoittautumiset:

(09) 4315 5333, kurssit@tieturi.fi, www.tieturi.fi



ELÄMÄ ON OPPIMISTA

POHJOISMAIDEN JOHTAVA ICT-KOULUTUKSEN JA PROJEKTIVALMENNUKSEN TARJOAJA