

**Organisaatio-orientoitunut tietovaraston kehitysprosessi
— tapaustutkimus tietovaraston suunnitteluprosessin hallinnasta
tietovarastointiprojektin toimittajaorganisaation näkökulmasta.**

Tommi Rantanen

Tampereen yliopisto
Tietojenkäsittelytieteiden laitos
Tietojenkäsittelyoppi
Pro gradu -tutkielma
Ohjaaja: Timo Niemi
Kesäkuu 2007

Tampereen yliopisto

Tietojenkäsittelytieteiden laitos

Tietojenkäsittelyoppi

RANTANEN TOMMI: Organisaatio-orientoitunut tietovaraston kehitysprosessi —
tapaustutkimus tietovaraston suunnitteluprosessin hallinnasta tietovarastointiprojektin
toimittajaorganisaation näkökulmasta.

Pro gradu -tutkielma, 102 sivua, 5 liitesivua

Kesäkuu 2007

Yritysmailmassa tietovarastoa tarvitaan sellaisten liiketoimintapäätösten tekemiseen, jotka perustuvat yrityksen tietojärjestelmien sisältämään informaatioon. Tietovarastolla on pitkäkestoinen rooli organisaation tietojärjestelmänä ja liiketoiminnan strategisen suunnan näyttäjänä. Tietovarastojen kehitysprojekteissa käytetyillä prosesseilla ja menetelmillä on merkittävä vaikutus suurien ja vaikutuksiltaan laaja-alaisten projektien hallintaan. Tämän tutkielman tarkoituksena on selvittää organisaatio-orientoituneen lähestymistavan merkitystä ja mahdollisuuksia tietovarastoinnille tietovaraston toimittajan perspektiivistä. Organisaatio-orientoituneen lähestymistavan lähtökohtana on kuvata asiakasorganisaation strategiset tavoitteet, jotka määrittelevät tietovaraston kehitysorganisaation tarpeet projektin tehokkaaseen elinkaaren ja metatietojen hallintaan. Keskeisin osa tätä tutkielmaa on organisaation tietomallin tuominen osaksi tietovarastointiratkaisua. Organisaation tietomalli perustuu asiakkaan liiketoiminnan tietojen käsitteelliseen mallintamiseen. Tutkielmassa suoritettiin tapaustutkimus, jonka kohteina oli kaksi hyvin erilaista tietovarastointiprojektia. Vastauksia tutkimukseen haettiin esimerkiksi kyselytutkimuksen avulla. Organisaation tavoitteiden asettaminen ja liiketoimintatietojen mallintaminen olivat ratkaisevassa asemassa, jotta tietovarastohankkeelle asetetut odotukset toteutuisivat. Tutkimuksen perusteella näyttäisi siltä, että yrityksissä ei toistaiseksi ole oivallettu tietovaraston keskeistä asemaa yrityksen tietoarkkitehtuurissa. Tietovarastointiympäristössä unohdetaan myös helposti perinteisten järjestelmäkehityksen ja tietovarastointiratkaisun eroavaisuudet ja tärkeät peruseriaatteen. On lisäksi huomioitava, että tietovaraston kehitysprojektissa metatietojen hallinta on merkittävässä roolissa. Tietovarastoratkaisujen tuotteistaminen saattaisi olla eräs ratkaisu liiketoiminta- ja tietokeskeisten tietovarastointiprojektien hallintaan.

Avainsanat ja -sanonnat: tietovarastointi, tietovarastointiprosessi, käsitteellinen mallintaminen, organisaation tietomalli, organisaatio-orientoitunut, tapaustutkimus.

Alkusanat

Pro gradu -tutkielman tekeminen oli erittäin haasteellinen ja pitkä prosessi. Mutta vaikka työ ja sen hallinta ajoittain tuntuikin vaikealta, oli mielenkiintoisen aiheen parissa työskentely kuitenkin myös hyvin antoisaa. Erityisesti työn aihepiirin läheinen liittyminen aikaisempiin työtehtäviini auttoi työn konkretisoimisessa ja samalla tutkimuksen tekeminen kartutti aihealueen käytännön osaamistani. Näin jälkikäteen on helppo todeta, että työn tekeminen kannatti ja oli ilo naulata filosofian maisterin tutkintooni viimeinen puuttuva naula.

Haluan kiittää työnantajaani Solitaa ja erityisesti teknologiajohtaja Jukka Jääheimoa mahdollisuudesta tehdä kirjoitustyötä osittain työajalla. Työni myös osittain edisti Solitan tietovaraston kehitysprosessien ja toimintatapojen parantumista. Matka asiakaslähtöiseen liiketoimintatietojen hallintaan on alkanut.

Kiitokset myös ohjaajalleni yliassistentti Timo Niemelle (FT), joka ymmärsi tutkielmani tekemisen päällimmäisen tavoitteen — valmistumisen. Hänen asiantuntevat ja tutkielmaa parantavat kommentit veivät tutkielmaa aina muutaman askeleen kohti tasokkaampaa lopputulosta.

Sisällys

1.	Johdanto	1
1.1.	Tutkimuksen tausta	2
1.1.1.	Aikaisempi tutkimus	3
1.2.	Tutkimuksen toteutus	4
1.2.1.	Tutkimuksen tavoite ja rajaukset	5
1.2.2.	Tutkimusmenetelmät ja -materiaali	6
1.3.	Tutkielman rakenne	9
2.	Tietovarastointi	9
2.1.	Mitä tietovarastointi on?	10
2.2.	Vaikutus ja arvo	13
2.3.	Tietovarastoinnin osa-alueet	16
2.4.	Tietovarastoarkkitehtuurit	19
3.	Tietovaraston tietomallit ja mallintaminen	20
3.1.	Tietovaraston tietomallit	23
3.2.	Käsitteellinen mallintaminen	25
3.2.1.	ER-mallintaminen	26
3.2.2.	Dimensionaalinen mallintaminen	26
3.2.3.	Muita malleja	29
3.3.	Looginen mallintaminen	30
3.4.	Mallintamisen vaikutuksia tietovaraston arkkitehtuuriin	32
3.5.	Organisaation tieto ja tietomalli	34
3.5.1.	Organisaation ei-tekniinen arkkitehtuuri	38
3.5.2.	Tietointegraatio	39
3.5.3.	Yrityksen tietoarkkitehtuuri	41
3.5.4.	Zachman Framework	43
3.5.5.	Tietomallin hyödyt	44
4.	Tietovaraston kehittämisprosessi	46
4.1.	Menestystekijät tietovaraston luomiseen	46
4.1.1.	Tietovarastoprosessin laadulliset tavoitteet	48
4.1.2.	Tietovaraston laadun mittarit	50
4.2.	Tietovaraston kehitysprosessi	50
4.2.1.	Tietovarastointiprosessin sidosryhmät	56
4.2.2.	Organisaatiolähtöinen kehittämisprosessi	57
4.2.3.	Suunnittelumallit	62
4.2.4.	Tietovaraston jatkuva kehittäminen	62
4.2.5.	Metatieto ja tietovarastointi	64

4.3.	Tietovaraston kehitysmetodit	65
4.3.1.	Tietovaraston kehittämisen erityispiirteitä	65
4.3.2.	Lähestymistapoja tietovaraston kehittämiseen.....	66
4.3.3.	Tietovaraston kehitysmenetelmiä.....	67
5.	Tapaustutkimukset.....	70
5.1.	Tapausten valinta ja esittely	71
5.2.	Materiaali ja tiedonkeruumenetelmät.....	73
5.3.	Tutkimusasetelma ja analyysiyksiköt	75
5.4.	Tapaus 1: yritys A	77
5.4.1.	Tietovarastointiprojektin laatutavoitteet	79
5.4.2.	Tietoarkkitehtuuri	79
5.4.3.	Tietovaraston kehitysprosessi	79
5.4.4.	Metatiedot, standardit ja ohjeet.....	80
5.4.5.	Tietovaraston määrittely ja kehittäminen.....	81
5.4.6.	Tietovaraston suunnittelumallit.....	81
5.4.7.	Tietovaraston tietomallit	82
5.4.8.	Muita huomioita.....	82
5.4.9.	Organisaatio-orientoitunut lähestymistapa ja organisaation tietomallit tapauksessa 1.....	83
5.5.	Lähtökohdat tapaustutkimukselle 2.....	83
5.6.	Tapaus 2: yritys B.....	83
5.6.1.	Tietovaraston tavoitteiden asettaminen.....	84
5.6.2.	Tietoarkkitehtuuri	84
5.6.3.	Tietovaraston kehitysprosessi	85
5.6.4.	Tietovaraston määrittely ja kehittämisen	87
5.6.5.	Tietovaraston tieto- ja suunnittelumallit	88
5.6.6.	Metatiedot, standardit ja ohjeet.....	88
5.6.7.	Vastaajien yleisiä kommentteja tietovarastoinnista	90
5.6.8.	Organisaatio-orientoitunut lähestymistapa ja organisaation tietomallit tapauksessa 2.....	91
6.	Tulokset ja havainnot.....	91
6.1.	Päätelmät tapaustutkimuksien osalta	92
6.2.	Tutkimuksen arviointi, luotettavuus ja jatkotutkimustarpeet	92
6.3.	Keskustelu	93
	Viiteluettelo	95
	Liitteet.....	103
	Liite 1. Kyselylomake tapaukselle 2.....	103

1. Johdanto

Tietovarastoinnin juuret ulottuvat kauas tietotekniseen historiaan. Tietotekniikkaa on käytetty liiketoiminnan päätöksenteon tueksi jo hyvin varhain, kun ymmärrettiin, että tekniikka voitiin valjastaa liiketoiminnan avuksi. 1970-luvun alussa rakennettiin ensimmäiset järjestelmät päätöksenteon tueksi. Näitä kutsuttiin *päätöksenteon tukijärjestelmiksi* (engl. Decision Support Systems, DSS) [Turban, 1993, s. 12]. Tämä on myös varhainen nimitys tietovarastoinnille. Tietokantoihin lisättiin *dimensionaalinen kehys* (engl. On-line analytical processing, OLAP-tuki), jotta niistä saatiin sopivampia liiketoiminnan moniulotteiseen analysointiin ja tiedon keruuseen. OLAP-tuen myötävaikutuksella ensimmäisten päätöksenteon tukijärjestelmien jälkeen syntyivät erityisesti yritysten johtajille kehitetyt *informaation hallintajärjestelmät* (engl. Executive Information Systems, EIS) [Turban, 1993, s. 394].

1990-luvun alkupuolella Bill Inmon ja Ralph Kimball tekivät tunnetuksi ideologian relaatiotietokantojen päälle rakennettavista päätöksenteontukijärjestelmistä. Tähän asti vallalla olivat olleet erilaiset suorkoneympäristöt. Kun dimensionaalinen mallintaminen kehittyi ja se yhdistettiin tiiviimmin EIS-järjestelmiin, syntyi käsite tietovarastointi [Arnott and Pervan, 2005]. Tästä eteenpäin kehitys on ollut nopeaa. On tapahtunut myös merkittäviä muutoksia teknisestä suuntautumisesta liiketoimintaprosessien tukemiseen. Tätä kuvaa hyvin nykyinen termi *liiketoimintatiedon hallinta* tai *älykäs liiketoiminta* (engl. Business Intelligence, BI). Se on yleispätevä termi, joka kuvaa organisaation sisäisen ja ulkoisen informaation huomioonottamista parempien liiketoimintapäätöksien saavuttamiseksi [Hannula and Pirttimäki, 2004]. Tämän termin alle sijoittuu tällä hetkellä monen eri alueen tietoteknistä kehitystä. Se ulottuu tiedon ja prosessien integroinnista tietovarastoinnin kautta raportointiin ja mittaristoihin.

Tällä hetkellä kaikki suuret ohjelmistotalot tarjoavat omia lähestymistapoja ja ratkaisuja päätöksenteon tueksi. Liiketoimintatiedon hallinta on tärkeä liiketoiminnan tukialue ja jokainen yritys, joka haluaa menestyä liiketoiminnassaan, on rakentanut tähän tarkoitukseen oman ratkaisunsa. Tietoteknisen uutisoinnin ja tutkimusyhtiöiden tutkimusten mukaan liiketoimintatiedon hallinta on tällä hetkellä yksi nopeimmin kasvavia tietotekniikan liiketoiminta-alueita.

Tietovarastointi on eräs keskeinen osa liiketoimintatiedon hallintaa. Vaikka tietovarastointi onkin tietoteknisessä mielessä vanha ja tutkittu aihe, sen kehitys on jatkuvaa ja sen pitää vastata jatkuvasti liiketoiminnan uusimpiin haasteisiin. Tämän tutkimuksen tarkoituksena on selvittää asiakasorganisaation liiketoimintatietojen vaikutusta tietovaraston kehitysohjelman suunnitteluvaiheisiin. Lisäksi tutkimus perehtyy tietovaraston kehittämisessä tarvittavien tietojen keruuseen, hallintaan ja soveltamiseen toimittajaorganisaation näkökulmasta.

Omalta kohdaltani kiinnostus tähän aiheeseen on syntynyt työni yhteydessä, koska olen havainnut, että tietovarastointitoteutusten taustalla prosessit, suunnittelumallit ja tiedonhallinta ovat oleellisia tekijöitä onnistuneessa tietovarastointiprojektissa. Soveltuvien kehitys- ja mallintamismenetelmien käyttäminen tietovarastointikehityksen tukena auttaa mielestäni luomaan tehokkaamman, pitkäkestoisemman sekä paremmin liiketoimintaa tukevan ratkaisun.

1.1. Tutkimuksen tausta

Tietovarastointiratkaisun toteuttaminen on samanlainen tietojärjestelmäprojekti kuin mikä tahansa muukin järjestelmä- tai ohjelmistoprojekti. Jotta lopputuloksena on onnistunut toteutus, kunnossa pitää olla samat osa-alueet kuin muissakin tietotekniikkaa hyväksikäyttävissä ratkaisuissa: projektinhallinta, resursointi, prosessit, määrittely, loppukäyttäjien kuuleminen, jne. [Adelman and Moss, 2000, s. 96]. Adelman ja Moss [2000, s. 11-13] toteavat, että tietovarastointiprojektin projektinhallinta eroaa kuitenkin perinteisistä projektinhallintatekniikoista esimerkiksi jatkuvasti muuttuvien vaatimusten, metatietojen tärkeyden, intensiivisen osallistumisen ja käyttäjälähtöisen suunnittelun takia. Mitä suurempi ratkaisu ja toteutus, sitä merkittävämmässä roolissa ovat prosessit, joilla projektia hallitaan.

Yksi tietovaraston kehittämisprosessin osaprosessi on liiketoiminnan käsitteiden ja niiden hallintaa avustavien tietomallien määrittely. Näiden merkitys on suuri, jotta kaikki projektin osallistujat ymmärtävät millaisesta maailmasta on kyse ja siitä olisi yhtenäinen käsitys. Liiketoimintatiedon hallinnassa ja erityisesti tietovarastoinnissa on pyrittävä yhdenmukaiseen ja yleispätevään toimintatapaan, jotta saavutetaan suurimmat liiketoimintaa tukevat hyödyt. Samaa mallia ja käsitettä voidaan hyödyntää tietokantaobjektin, yrityksen sovellusintegraation (engl. Enterprise Application Integration, EAI) tietomallien tai liiketoimintaprosessin määrittämisessä. Sama malli on voitava jakaa teknisen järjestelmän kehittäjien ja liiketoimintaa tuntevien asiantuntijoiden sekä analyytikoiden kesken. Tällöin yhteinen näkökulma on helpommin määritettävissä.

Valmiita malleja on olemassa useita, vaikka suurimmalla osalla niistä ei välttämättä ole muuta eroa kuin nimi. Suurilla tietotekniikkataloilla on omia kaupallisia tai avoimia malleja ja lisäksi on tietotekniikan tutkijoiden luomia malleja. Näitä ovat mm. UDM (Microsoftin Unified Dimensional Model) [Sanders, 2005] ja erilaiset yleiskäyttöiset tietomallit (engl. universal data model), jollainen on esimerkiksi Oraclen Customer Data Hub [Oracle, 2006]. Mallien laajuudessa ja soveltuvuudessa on kuitenkin eroja. Malli voi olla esimerkiksi valmis ratkaisu suuryrityksen tarpeisiin. Tässä tutkimuksessa keskityn tutkimaan yrityskohtaisen ja kokonaisvaltaisen suunnittelutason tietovaraston kehitysmallin hyödyntämismahdollisuuksia. Tutkielman keskeisenä osana arvioin kokonaisvaltaisen organisaation tietomallin käytöllä saavutettavia hyötyjä tietovarastoinnin kehitysprosessin osana. Yrityksen tietomalli on organisaation

tietomalli, jota voidaan hyödyntää kohdeorganisaatiossa kaikkialla yhtiön toimialueella. On huomioitava, että käsittelemäni organisaation tietomalli on laajempi kokonaisuus kuin kirjallisuudessa tähän asti kuvattu organisaation tietomalli tai perinteinen käsitteellinen tietomalli. Lisäksi esittelemäni organisaation tietomalli eroaa yleiskäyttöisistä tietomalleista, jotka ovat monen yrityksen liiketoimintaan sopivia yleispätevällä tavalla esitettyjä tietomallikokoelmia [Silverston, 2001, s. 2-3].

Tutkimukseni kohteena olevan organisaation tietomalli sisältää tietomallin lisäksi esimerkiksi organisaatiosidonnaisia toimintatapoja. Kokonaisvaltainen tietovarastoinnin kehittäminen on esitelty tutkimuksessani *organisaatio-orientoituneena lähestymistapana*, jossa organisaation erilaiset kehitysmenetelmä-, mallintamis- ja dokumentointitarpeiden huomioon ottaminen toimii onnistuneen, yhteiseksi hankkeeksi otetun ja kestävän tietovarastoinnin perustana. Tietovaraston kehitysprojekteihin kehittämäni lähestymistapa yhdistää *asiakasorganisaation tavoitteiden asettamisen ja tietovarastoa kehittävän organisaation kehitystarpeet hyödyntämällä organisaation tietomallia ja ottamalla huomioon tietovaraston kehitystä tukevien meta- eli liitännäistietojen merkityksen*.

Tutkielmassani keskityn liittämään yrityksen tietomallin kehittämisen osaksi tietovarastointihankkeen suunnitteluprosessia. Mallien soveltaminen kehitysprosessin osana mahdollistaa useiden lähestymistapojen käyttämisen kuten esimerkiksi osittavan (”ylhäältä-alas”) (engl. top-down) ja kokoavan (”alhaalta-ylös”) (engl. bottom-up) lähestymistavan. On myös kiinnitettävä huomiota siihen, että suunnitteluprosessilla ja yrityksen tietomallilla on kiinteä vuorovaikutus koko hankkeen ajan. Tutkin myös tietovaraston kehitysprosessin kokonaisuuden hallinnalle merkityksellisen tietovarastojen metatietojen hallinnan vaikutuksia tietovarastohankkeen onnistumiselle ja tehokkuudelle.

1.1.1. Aikaisempi tutkimus

Näyttää siltä, että aikaisempia tutkimuksia yrityksen tietomalleista, organisaatio-orientoituneiden tietomallien käytöstä tai soveltamisesta tietovarastoinnissa on tehty hyvin vähän. Tietovarastoinnin näkökulmasta aihetta on sivuttu pääasiassa vain loogisen mallintamisen osalta, mutta aihetta on kuitenkin tärkeä tutkia, koska asiakas hyötyy kokonaisvaltaisemmasta näkemyksestä ja mahdollistaa asiakkaan liiketoiminnan tavoitteellisemman kehittämisen. Toisaalta eri toimialoille ja organisaatioille on luotu omia tietomalleja ja esimerkiksi Yhdysvaltojen terveystieteiden ministeriön (U.S. Department of Health & Human Services) hallinnon alainen palvelukeskus (Centers for Medicare & Medicaid Services) on laatinut ohjeelliset suositukset, suunnitteluprosessi- ja menetelmäkuvaukset sekä tietostandardit ministeriön tietovarastointihankkeille [CMS, 2007]. Nämä ohjeet esittelevät esimerkiksi tietovaraston mallintamiseen standardin, jonka avulla organisaation tietomalli luodaan ja miten sitä hyödynnetään tietovarastointihankkeissa. Vähäisestä tutkimuksesta päättelin, että yrityksen tietomalli

on kaupallisten toimijoiden kehittämä malli. Joten käsitteellinen liiketoimintalähtöinen mallintaminen yhdistettynä tietovarastointiin on näkökulmaltaan tai aiheeltaan sellainen, josta tutkijat eivät välttämättä ole kiinnostuneita. Aihe voi myös olla liian monisyinen konkreettiseen tutkimukseen. Kolmas vaihtoehto on, että tutkijat ovat käyttäneet aiheesta eri termejä kuin niitä, joita minä aihetta tutkiessani määrittelin.

Kokonaisvaltaista tietovaraston kehittämisprosessia on myös tutkittu melko vähän. Alan perusteoksissa [esimerkiksi Inmon, 2005; Ponniah, 2001; Kimball and Ross, 2002] kehitysprosessia on yleensä kuitenkin kuvattu jollakin tapaa. Pereira ja Becker [2000] ovat vertailleet erilaisia tietovaraston kehitysmetologioita ja he ovat esitelleet kokeiluprojektiin perustuvan tietovaraston kehitysmenetelmän erityisesti sellaisten yritysten tarpeisiin, joilla ei ole aikaisempaa osaamista tietovarastoista.

Stefanov ja List ovat tutkineet [2006] liiketoimintatietojen ja organisaation tavoitteiden vaikutusta tietovarastoinnille. He esittelevät erään mallin liiketoiminnan tavoitteiden ja tietovaraston tietomallin yhdistämiseksi. He jatkavat aiheesta tutkimuksessaan [2007], joka käsittelee organisaation mallin ja siitä johdetun tietämyksen linkittämistä tietovarastoon. Lähtökohtana tutkimuksessa on mallintaa liiketoimintatietoa tietovaraston päälle ja muodostaa uusi abstraktiokerros liiketoimintaihmissen käytettäväksi. Stefanov ja muut [Stefanov et al., 2005] ovat myös tutkineet liiketoimintaprosessien ja käsitteellisten mallien suhdetta tietovarastoympäristössä.

Vassiliadis [2000] on tarkastellut vuosien 1995 – 1999 välillä julkaistuja tietovarastointiin liittyviä tutkimuksia ja hänen havaintojensa perusteella voi todeta tutkijoiden keskittyvän pääasiassa teknisiin ongelmiin. Nämä 99 tutkimusta keskittyivät epätäydelliseen informaatioon, integrointiin, OLAP-mallintamiseen, kyselyiden käsittelyyn, päällekkäisen tiedon käyttöön, tallennuksen hallintaan ja näkymäteknologiaan. Lisäksi kirjallisuuskartoitusta tehdessäni havaitsin, että tutkijoita ovat erityisesti kiinnostaneet tietovaraston laatuasiat sekä hankkeiden onnistumiseen liittyvät elementit.

Herrmann [2004] on todennut tutkimuksessaan saman asian kuin Vassiliadiksen [2000] tutkimus osoittaa. Ammatillaiset ja tutkijat keskittyvät tietovarastoinnin alueella tekniseen tarkastelukulmaan. Herrmann on tutkinut joitakin haasteita sekä ongelmia organisatorisesta näkökulmasta huomioiden myös tietovarastointiratkaisun kehitysprosessin hallinnan.

1.2. Tutkimuksen toteutus

Tutkielma lähestyy tietovarastokehityksen ongelmakenttää tapaustutkimuksen (engl. case study) näkökulmasta ja erilaisia aineistoja hyväksikäyttäen. Yrityksessä X, jossa työskentelen, tietovaraston suunnittelijat ja kehittäjät ovat havainneet tarpeelliseksi parantaa tietovarastointiprojektien läpiviennin kokonaisvaltaista hallittavuutta.

Organisaatio-orientoitunut lähestymistapa on eräs tutkimuksen keskeinen käsite, jolla kuvaan kokonaisvaltaista, organisaation tiedon hallintaan perustuvaa tapaa toteuttaa tietovaraston kehitysprojekti. Se sisältää organisaation kattavien tietomallien hyväksikäytön ja toimittajaorganisaation suunnittelumenetelmät sekä kehitysprosessien hallitsemisen organisaation nykyisten ja tulevien tarpeiden perusteella sivuuttamalla tekniset näkökohdat.

Tutkimuksen teoreettisen viitekehyksen on tarkoitus tukea edellä mainittua lähestymistapaa. Tämän tutkimuksen teoreettinen viitekehys koostuukin seitsemästä keskeisestä käsitteestä, joista jokainen esitellään tarkemmin luvuissa 2-4. Nämä keskeiset käsitteet muodostavat tietovarastoprojektin kronologisen suunnittelumenetelmän esimerkiksi seuraavalla tavalla: aluksi määritellään tietovarastointiprojektin *laatutavoitteet* (1). Tämän jälkeen muodostetaan tietovaraston tietosisältö määrittämällä organisaation *tietoarkkitehtuuri* (2). Jotta tietovarastoa voidaan hallitusti viedä eteenpäin, tarvitaan tietovaraston *kehitysprosessi* (3). Pohjaksi kaikelle suunnittelulle ja määrittelylle tarvitaan yhteisesti sovitut *standardit* ja *ohjeet* (4). Seuraavaksi suoritetaan tietovaraston *määrittely metatietoja* (kuten tietoa liiketoimintaprosesseista) keräämällä ja sopivia *kehitysmenetelmiä* hyväksikäyttäen (5). Tietovaraston suunnitteluun tarvitaan *suunnittelumalleja* (6). Kun tietovaraston perusta on kunnossa, organisaation tieto kuvataan *tietomalleilla* (7). Nämä analysointiyksiköt esitellään tapaustutkimuksen kontekstissa tarkemmin luvussa 5.

1.2.1. Tutkimuksen tavoite ja rajaukset

Tutkimuksen keskeinen perusrakenneosa on organisaatiossa oleva tieto. Organisaation tietoja on tulkittava, mallinnettava ja organisoitava, jotta niitä voidaan hyödyntää tietovaraston suunnittelussa sekä kehitysprosessissa. Näillä toimenpiteillä muodostan tietovaraston kehitysprojektille kehyksen, jonka avulla mahdollistetaan organisaatiota tukevien tietojen parempi hallinta.

Tutkimuksen tavoitteena on selvittää, miten organisaation liiketoimintatiedon hallintaa voidaan parantaa *organisaation tietomalliin* (engl. Corporate data model / Enterprise Data Model, EDM) perustuvan tietovarastointiratkaisun avulla sekä miten organisaatio-orientoituneen lähestymistavan käyttämisellä saavutetaan parempi tietovaraston kehitysprojektin hallittavuus. Edellä esittelemäni organisaation tietomalli sisältää formaalin käsitteellisen tietomallin lisäksi laajemman näkökulman tietovaraston tietojen ja käsitteiden määrittelyyn. Perinteinen formaali tietomalli on tietojen organisoinnin väline, kun taas organisaation tietomalli koostuu tämän lisäksi osittain esimerkiksi ohjeista, työjärjestyksestä, prosessikuvauksista ja menetelmätavoista, jotka sijoittavat sen tietovaraston kehitysprosessin osaksi.

Tämän tutkimuksen perusolettamukseni on se, että organisaation kattava tietomalli ja organisaatiota tukeva lähestymistapa auttavat saavuttamaan paremman lopputuloksen: ymmärrettävämmän, ylläpidettävämmän ja eri toimintasektorit ylittävän

tietovaraston. Tämä kokonaisvaltainen lähestymistapa tukee tietovaraston toteuttamisprojektia, tietovaraston kehitysprosessia sekä toimii lopulta välittäjänä tietovarastointiratkaisun integroinnissa muihin osajärjestelmiin ja tietoa soveltaviin järjestelmiin.

Tarkastelun kohteeksi olen valinnut yhden tietovaraston tärkeimmistä osa-alueista, suunnittelun. Sitä voidaan lähestyä esimerkiksi kysymällä, kauanko kuluu aikaa uuden vaatimuksen lisäämiseen järjestelmään? Mikä esittämässäni mallissa on parempaa kuin perinteisessä tietovaraston suunnittelussa ja toteutuksessa ja mikä mahdollisesti huonompaa? Miten kehittämisprosessia ja tietovaraston toteutusta voisi parantaa vielä edelleen?

Tutkielmani tutkimusongelman olen kiteyttänyt seuraaviin kysymyksiin:

- 1.) *mitkä ovat koko organisaation kattavan tietomallin hyödyt tietovaraston kehitysprosessissa ja liiketoiminnan tietojen kuvaamisessa?*
- 2.) *miten organisaatio-orientoitunut lähestymistapa voisi tukea tietovaraston kehitysprosessia?*

Tutkielmani ei ota kantaa tietovaraston projektinhallintaan. Vaikka tietovaraston projektinhallinta eroaakin tavallisesta ohjelmistoprojektin projektinhallinnasta, tietovarastoprojektin projektinhallintaa koskevat samat lainalaisuudet ja toimintatavat kuin tavallista ohjelmistoprojektin projektinhallintatyötä. Tutkielmassani ei myöskään oteta huomioon tavallisia projektin läpivientiin vahvasti vaikuttavia taustatekijöitä kuten projektihenkilöiden osaamistasoa tai käytettävissä olevia työvälineitä. Lisäksi projektien lähtötilannetta ei käsitellä, joten projektin ulkopuolelta tulevat vaatimukset sekä asiakasorganisaation peruspiirteet, kuten toimintatapa tai organisaatorakenne on rajattu tutkielman ulkopuolelle. Tämä johtuu siitä, että projektiryhmä tai sen jäsenet harvoin (tai hankalasti) voivat vaikuttaa organisaatioon, sen liiketoimintaan, organisaation asettamiin tavoitteisiin ja vaatimuksiin sekä käytettävissä oleviin resursseihin.

1.2.2. Tutkimusmenetelmät ja –materiaali

Tutkimukseni on jaettu kolmeen osaan, jotta käytetyt tutkimusmenetelmät tulisivat paremmin esille. Ensimmäisessä vaiheessa suoritan kirjallisuuskartoituksen organisaatio-orientoituneen kehitysprosessin osa-tekijöiden eli tapaustutkimukseni analysointiyksiköiden löytämiseksi. Tämän jälkeen pyrin konkretisoimaan havaintojani tapaustutkimusten avulla. Toisessa vaiheessa teorian perusteella muodostetut analysointiyksiköt toimivat ensimmäisen tapaustutkimuksen kohteen, yritys A:n viitekehyksenä. Tapausta kuvailemalla pyrin löytämään vastaavuuksia teoreettisen perustan osatekijöistä. Kolmannessa vaiheessa pyrin vahvistamaan ja selittämään tutkimustuloksia analysoimalla toisen tapauksen eli yrityksen B havaintoja.

Tutkimuksen voi toteutustavan aihepiirin ja tutkimusmahdollisuuksien perusteella sijoittaa moneen eri tutkimusmenetelmän kategoriaan. Järvinen ja Järvinen [2000, s. 9] ovat luoneet informaatioteknologian tutkimuksessa käytettäville tutkimusotteille taksonomian, jonka avulla esittelen käyttämäni menetelmän piirteitä. Lähtökohdiltaan tutkimukseni sijoittuu reaali maailmaa käsitteleviin tutkimusotteisiin. Nämä tutkimustyytit voidaan vielä erotella teorian käsittelytavan perusteella käsitteellisteoreettisiin, teoriaa testaaviin ja teoriaa luoviin [Järvinen ja Järvinen, 2000, s. 9].

Erikssonin ja Koistisen [2005, s. 4-5] mukaan useat menetelmäoppaat suosittelevat tapaustutkimuksen valitsemista lähestymistavaksi, jos tutkijalla on vähän kontrollia tapahtumiin ja aiheesta on vain vähän empiiristä tutkimusta ja tutkimuskohteena on jokin nykyhetken reaali maailman ilmiö. Lisäksi Galliers ja Land [1987] ovat tietojärjestelmätieteen tutkimuksen lähestymistapojen luokittelussaan sijoittaneet tapaustutkimuksen käyttökelpoiseksi lähestymistavaksi organisaatioiden ja metodologioiden tutkimiseen. Koska edellä mainitut piirteet sopivat tutkimukseeni, on tutkimusotteenani empiirinen tapaustutkimus. Tukea valintaani saan Darken ja muiden [1998] huomioista, että tapaustutkimus soveltuu erittäin hyvin informaatioteknologian innovaatioiden ja organisaatio-orientoituneen välisen vuorovaikutuksen ymmärtämiseen sekä tietojärjestelmien kehittämisen tutkimiseen. Tutkimus tukeutuu kvalitatiiviseen tutkimuksen edellyttämään tutkimustapaan. Siksi tutkimukseni sijoittuu kuvailevien, tulkitsevien ja uusia teorioita luovien lähestymistapojen luokkaan. Jos olisin mitannut ja arvioinut vertailemalla kahta eri tapausta, tietovaraston kehitysprosessia ilman yrityksen tietomallia ja sen kanssa, olisi kvantitatiivisen lähestymistavan käyttö ollut mahdollista. Kvalitatiivisen toteutustavan valintaa tukee Parén [2001] suorittama tutkimus, jossa hän on analysoinut erilaisia tapaustutkimustyyppejä. Hän on väittää, että heikkouksistaan huolimatta kvalitatiivinen tutkimustapa on paras lähestymistapa tutkittaessa monimutkaisia ilmiöitä, kuten järjestelmien suunnittelua ja toteutusta.

Tapaustutkimusta voidaan tutkia, suorittaa ja esittää monin eri tavoin ja siksi se voidaan määrittellä monin tavoin. Esimerkiksi Palvia ja muut [Palvia *et al.*, 2003] määrittelevät, että

“Tapaustutkimus on tutkimus organisaation yksittäisestä ilmiöstä
(esimerkiksi sovellus, teknologia tai päätös) lyhyehkön ajanjakson aikana.”

Cunningham [1997] on tunnistanut yhdeksän erilaista tutkimusmetodia tapaustutkimuksen suorittamiseksi. Ne voidaan jakaa kolmeen tyyppiin: vertailua korostaviin tutkimuksiin (engl. comparative case studies), intensiivisiin tutkimuksiin (engl. intensive case studies) sekä toimintatutkimuksiin (engl. action research cases). Näistä intensiiviset tapaustutkimuksen toteuttamistavat pyrkivät tulkitsemaan tutkittavan tapauksen, kuten tietyn organisaation tapahtumia ja käytäntöjä kuvailemalla ja tulkitsemalla tyypillisiä kokemuksia, joiden perusteella voidaan luoda uutta ymmärrystä ja teoriaa [Cunningham, 1997]. Eriksson ja Koistinen [2005, s. 15]

kuvaavat intensiivistä tapaustutkimusta kokonaisvaltaiseksi ja laadulliseksi tutkimukseksi, jossa tapausta tutkitaan tutkimuksen osallistujien näkökulmasta perustuen heidän omiin käsitteisiinsä ja kieleensä. Eriksson ja Koistinen korostavat, että tapausten lukumäärä vaikuttaa tutkimuksen orientaatioon. Koska tutkimuksessani on vain kaksi tapausta, on tutkimukseni täten selkeästi intensiivinen tapaustutkimus.

Intensiivisistä tapaustutkimuksista voidaan vielä erottaa esimerkiksi taulukointimetodi (engl. tabulation), selittävä (engl. explanatory) ja tulkitseva (engl. descriptive) tapaustutkimus [Cunningham, 1997]. Selittävä tapaustutkimus tuottaa tosiasioita ja näiden perusteella vedettäviä johtopäätöksiä lehtitoimittajien tapaan [Cunningham, 1997]. Edellä mainituille metodeille on yhteistä pyrkimys teorian testaamiseen [Järvinen ja Järvinen, 2000, s. 78]. Näiden lisäksi on monia muita tapaustutkimusmetodeja, kuten esimerkiksi kuvaileva tapaustutkimus. Sen tavoitteena voi olla esimerkiksi havaitun käytännön kuvailu arkipäivän tilanteessa. Tapaustutkimuksen suorittamisen piirteet ja käytetyt tapaukset on kuvattu luvussa 5.

Tapaustutkimuksen suunnitteleminen ja laajuuden määrittäminen vaatii perusteellisen kirjallisuuden analysoinnin, jotta tutkija pystyy ymmärtämään tutkimuskohteensa taustan ja yhdistämään se tutkimuskysymykseensä liittyvään kirjallisuuteen [Darke *et al.*, 1998]. Kirjallisuuskartoitus onkin tutkimusprosessini ensimmäinen vaihe. Varsinaisen tapausten käsittely on jaettu kahteen osaan. Ensimmäisen tapaustutkimukseni metodi on kuvaileva ja toisen metodi on selittävä. Näin pyrin saamaan monipuolisen kuvan tutkittavasta aiheesta. Johtuen käyttämästäni tiedonkeruumenetelmistä tapaustutkimukseni sisältää myös konstruktivisen tutkimuksen ja innovaation arvioinnin piirteitä.

Tutkimuksessani käyn läpi organisaatio-orientoituneen lähestymistavan piirteitä ja siihen keskeisesti liittyvän organisaation tietomallin piirteiden toteutumista operatiivisessa käytössä olevan tietovaraston kehitysprojektissa. Vertailukohtena käytän tietovarastointiprojektia, jossa kehitystyö aloitettiin ilman aikaisemman tietovaraston olemassaoloa ja sen aiheuttamien valintojen painolastia.

Materiaalina käytän tietovaraston kehitysprojektien dokumentaatiota sekä dokumentoimattomia vallitsevia käytäntöjä ja tapoja toimia. Ensimmäisen tapauksen osalta käytän osallistuvaa havainnointia ja arviointia. Lisäksi ensimmäisessä tapaustutkimuksessa voidaan käyttää projektin eri vaiheiden sähköpostikommunikaatiota. Myös projektien käytävissä olevat tietojen hallintaan liittyvät työkalut ja niiden käyttötapa prosessin osana sisältyvät arviointiini. Toinen tapaus on luonteeltaan selittävä, joten tätä vahvistavaa vertaistapausta tutkin suullisen ja kirjallisesti annetun palautteen perusteella. Projektin asiantuntijoiden palaute on kerätty projektin aikana ja sen jälkeen. Kirjallisesti annettu palaute on ollut sekä vapaamuotoista että strukturoitua perustuen kyselylomakkeeseen. Lisäksi käytössäni on

projektissa tuotettu dokumentaatio. Tarkemmin materiaali ja tapaukset on kuvattu luvussa 5.

1.3. Tutkielman rakenne

Tutkielma on jaettu kuuteen lukuun. Ensimmäinen luku sisältää johdannon tutkimuksen aiheeseen ja keskeisiin käsitteisiin sekä esittelee käytetyt tutkimusmenetelmät. Toisessa luvussa esitellään tutkimukseni kannalta oleellisia tietovarastoinnin keskeisiä peruseriaatteita. Kolmannessa luvussa tutkin tietovarastoitavaa tietoa, tietomalleja sekä tiedon mallintamista ja hallitsemista. Tämä luku muodostaa organisaatio-orientoituneen lähestymistavan tietokeskeisen ytimen. Neljäs luku käsittelee tietovaraston kehittämis- ja suunnitteluprosesseja ja laajennan organisaatio-orientoituneen lähestymistavan koskemaan koko tietovaraston suunnitteluprosessia. Viidennessä luvussa käyn läpi suorittamani tapaustutkimukset, joiden tarkoitus on muodostaa yhdessä teoreettisen viitekehyksen kanssa uutta teoriaa ja selityksiä organisaatio-orientoituneen lähestymistavan tueksi. Viimeisessä luvussa käyn läpi tutkimuksessa löydetty keskeiset havainnot sekä esittelen mahdollisia jatkotutkimusaiheita.

2. Tietovarastointi

Harding ja Yu [1999] esittelevät tietovarastoinnin periaatteena, joka tarjoaa organisaation käyttäjille integroidun ja keskitetyn informaatiolähteen, jota voidaan hyödyntää analyttisesti tai informaation kannalta tuottamaan johdolle raportteja heidän päätöksentekonsa tueksi. Teknisessä mielessä ja yksinkertaistettuna tietovarastointi perustuu tiedonkeruuseen ja sen jalostamiseen eli yhteenvetoinformaation muodostamiseen tietokannoissa ja raporteissa olevasta informaatiosta.

Organisaation informaatio on yksi sen tärkeimmistä arvotekijöistä ja kilpailuvalteista. Tietovaraston rooli informaation hallinnassa on jalostaa tietoa ulospäin ja operatiivisen järjestelmän rooli on puolestaan kerätä tietoa sisäänpäin. Organisaation informaation hallinnan näkökulmasta katsottuna operatiivisen järjestelmän käyttäjät pyörittävät organisaation rutiineita, kuten ottavat vastaan tilauksia ja kirjaavat asiakkaiden valituksia. Tietovaraston käyttäjät puolestaan seuraavat organisaation rutiinien kehittymistä, kuten vertaavat uusien tilauksien määriä edellisen viikon määriin tai tutkivat mistä seikoista asiakkaat valittivat. [Kimball and Ross, 2002, s. 2]

Inmon [2005, s. 20] on määritellyt tietovaraston ja päätöksentekojärjestelmän käyttäjät analyttikoiksi, liiketoimintaihmisiksi, joilla on myös teknistä osaamista [Inmon, 2005, s. 20]. Toisaalta Ponniah [2001, s. 323] on todennut, että jokainen, joka tarvitsee strategista informaatiota, on tietovaraston käyttäjä. Näitä käyttäjiä ovat esimerkiksi liiketoiminta-analyttikot, liiketoiminnan suunnittelijat, osastopäälliköt ja yrityksen johtajat. Inmon [2005, s. 20] on esittänyt analyttikon tärkeimmäksi tehtäväksi määrittellä ja löytää informaatiota, jota käytetään yrityksen päätöksenteossa. Ponniah

[2001, s. 323-325] on luokitellut käyttäjät heidän tarvitsemansa informaation, tietoteknisten käyttötaitojen, työtehtävien sekä organisaatiotasojen mukaan viiteen ryhmään:

- Kaivajiin (engl. miners), jotka ovat erikoistuneita analyytikkoja ja kiinnostuneita tiedon löytämisestä.
- Tutkijoihin (engl. explorers), jotka ovat taitavia analyytikkoja ja kiinnostuneita ennalta ennustamattomista analysointitarpeista (engl. ad hoc analysis).
- Kasvattajiin (engl. farmers), jotka ovat analyytikkoja ja kiinnostuneita ennalta tiedetyistä analysointitarpeista.
- Turisteihin (engl. tourists), jotka ovat johtajia ja kiinnostuneita liiketoiminnan mittareista.
- Operaattoreihin (engl. operators), jotka ovat tukihenkilöstöä ja kiinnostuneita nykyisistä tiedoista.

Inmon [2005, s. 20] on todennut, että usein analyytikot ympäri maailman määrittelevät vasta saamiensa tietojen perusteella, mitä he oikeastaan haluaisivat nähdä. Tämä käytäntö johtuu analyytikoiden tavasta hallita liiketoimintaansa. Myös Gardner [1998] on muistuttanut, että tietovaraston käyttäjien nykyiset kysymykset eivät välttämättä ole samoja huomenna. Lisäksi olennaista on havaita, että tällaisella ajattelutavalla on merkittävä vaikutus tietovaraston kehittämisprosessiin [Inmon, 2005, s. 20].

2.1. Mitä tietovarastointi on?

Tietovarastoinnin voi määritellä eri tavoin näkökulmasta riippuen. Tässä tutkielmassa käyttämäni liiketoiminnallinen näkökulma tulee ilmi esimerkiksi Gardnerin [1998] määrittämisestä:

”Tietovarastointi on prosessi - ei tuote - tietojen kokoamiseen ja hallinnointiin erilaisista tietolähteistä, jotta saavutetaan yksi, yleinen näkymä osaan tai kaikkeen liiketoimintaan.”

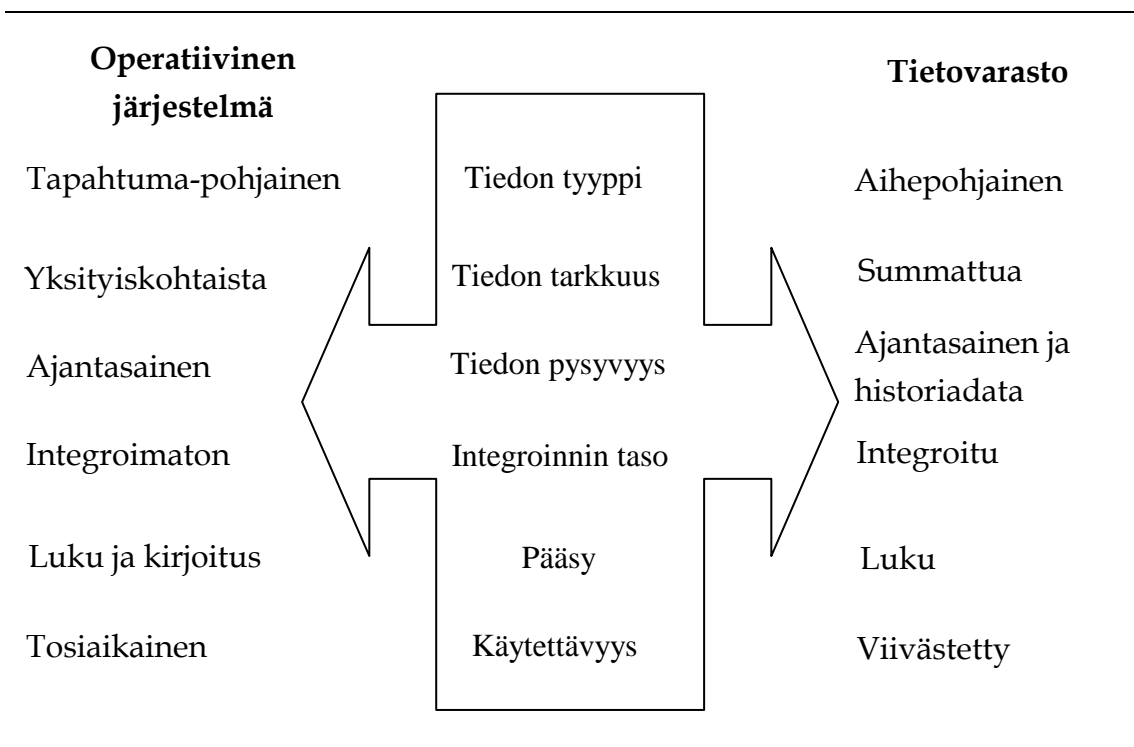
Erilaiset tietolähteet ovat yleensä operationaalisia tietokantoja, mutta ne voivat olla myös tiedostoja, yrityksen ulkopuolisia lähteitä, joista tieto kerätään esimerkiksi integraation keinoin.

Tietovarastoinnin peruskomponentti on tietovarasto (engl. data warehouse). Mielestäni tietovarasto on liiketoimintatiedon hallinnan keskeisin osa-alue, koska tietovarastoon pyritään keräämään kaikki päätöksenteon kannalta oleelliset tiedot. Lisäksi tietovarasto on yleensä yrityksen ainoa resurssi, josta käy ilmi liiketoiminnasta kerätyn tiedon historiallinen perspektiivi. Tietovarasto onkin yleensä aina tavalla tai toisella mukana liiketoimintatiedon hallinnassa [Hannula and Pirttimäki, 2004]. Inmon [2002, s. 389] määrittelee tietovaraston teknisen näkökulman seuraavasti:

Tietovarasto on kokoelma integroituja, aihekohtaisia (engl. subject-oriented) tietokantoja, jotka ovat suunniteltu tukemaan päätöksentekoa, ja joiden jokainen tietoyksikkö on suhteessa johonkin ajan hetkeen. Tietovarasto sisältää perustietoa ja niihin liittyvää summattua yhteenvetotietoa.

Kimball ja Ross [2002, s. 397] puolestaan korostavat, että tietovarasto sisältää operatiivista järjestelmistä (liiketoiminnan kannalta kriittisiä järjestelmiä) kerättyjä tietoja, jotka ovat organisoitu yhteenvetokyselyjen ja analysoinnin tehokkuuden ja helppokäyttöisyyden näkökulmasta. Liiketoimintaa varten luotujen tietovarastojen avulla pyritäänkin yleensä havaitsemaan liiketoiminnan trendejä, esimerkiksi tietyn tuotteen myynnin kehitystä ja kannattavuutta.

Tietovarastot eroavat operatiivisista järjestelmistä monessa suhteessa. Winter [2001] esittelee näitä eroja tietokantaan talletettujen tietojen näkökulmasta kuvassa 1.



Kuva 1. Operatiivisen järjestelmän ja tietovaraston piirteiden merkittävimmät erot talletetuissa tiedoissa ja niiden hyödyntämisessä.

Näiden kahden tietojärjestelmätyypin sisältämien tietojen ominaispiirteitä on kuvassa 1 kuvattu kuuden attribuutin suhteen. Tietovarasto on yleensä mallinnettu liiketoimintalueen aihesisällön suhteen, kun operatiivisen järjestelmän pääpaino on tapahtumapohjaisessa tietojenkäsittelyssä. Tietovarasto säilyttää tämänhetkisen tiedon lisäksi myös historiatietoa muutosten ja trendien analysoimiseksi. Tietovaraston tiedot eivät yleensä ole kuitenkaan reaaliaikaisia kuten operatiivisen järjestelmän sisältämät tiedot, vaan tietojen päivitys suoritetaan massapäivityksinä ja sopivin väliajoin käytetystä aikahierarkiasta riippuen esimerkiksi kerran päivässä. Tietovaraston tietojen pääasiallisina käyttötapoina ovat yksinkertaiset kyselyt ja päivityksiä tehdään vain

tarvittaessa, mutta säännöllisesti, kun taas operatiivisen järjestelmän dataa päivitetään ja luetaan jatkuvasti. Tietovaraston tiedot ovat yleensä operatiivisen järjestelmän tiedoista muodostettua yhteenvedotietoa, joka usein edellyttää integrointia eri operatiivisten järjestelmien kesken. Tiettyä operatiivista järjestelmää ei ole läheskään aina integroitu toisiin operatiivisiin järjestelmiin.

Thomsenin [1997, s. 9] erittelyanalyysissä tietojen prosessointinäkökulmasta katsottuna on lisäksi eroja käsittelyoperaatioiden ennustettavuudessa, kyselyissä tarvittavassa tiedon määrässä ja yksityiskohtaisuudessa sekä tiedon johdettavuudessa. Operatiivisten järjestelmien toiminnot ovat ennustettavia ja kyselyt kohdistuvat pieneen määrään talletettuja tietoja eli ne ovat harvoin laskennallisesti muodostettu toisesta tiedosta. Tietovarastoon pohjautuvat analysointitarpeet ovat huomattavasti ennustettavia ja ne perustuvat suurempaan määrään talletettuja tietoja, jotka ovat jalostettua ja mahdollisesti monimutkaisesti johdettu lähtötiedoista [Thomsen, 1997, s. 9]. Tällaisia johdettuja tietoja ovat esimerkiksi tietyt tilastolliset suureet kuten varianssi ja keskihajonta.

Näiden kahden erityyppisen tietojärjestelmän käsittelyä kuvaamaan on vakiintunut kaksi termiä. Operatiivisista järjestelmistä käytetään lyhennettä OLTP (Online Transaction Processing) ja tietovaraston analysointijärjestelmistä OLAP (On Line Analytical Processing) [Thomsen, 1997; Inmon et al., 1997; Halpin, 2001]. Halpin [2001, s. 708] määrittää OLTP-järjestelmät tyypillisiksi tietokantaperustaisiksi sovelluksiksi, joihin liittyvät useasti toistuvat tiedon päivitykset ja kyselyt.

OLAP on tietovaraston sisältämän yhteenvedotiedon käsittelyä — toiminnallisuutta, jonka tarkoituksena on helpottaa tietojen moniulotteista analysointia [Inmon et al., 1997, s. 183-184]. Myös Shoshani [1997] sisällyttää OLAP:n tietovaraston osaksi, koska OLAP-toiminnallisuus perustuu liiketoimintasovellusten tuottaman tiedon analysointiin, jonka tarkoituksena on avustaa tietyn tyyppistä päätöksentekoa. Päätöksentekijät ovat yleisesti tietyn liiketoiminta-alueen johtajia [Shoshani, 1997]. Jo toimintansa lopettanut OLAP Council-organisaatio puolestaan määrittelee OLAP-termin ohjelmistoteknologiana, joka auttaa analyytikkoja ja johtajia saavuttamaan nopeasti ja interaktiivisesti yhdenmukaisen ja integroidun käsityksen liiketoimintojen tilasta. Samalla se tarjoaa pääsyn tietoon, joka on muodostettu talletetuista tiedoista heijastaen yrityksen keskeisiä tekijöitä – toisin sanoen ulottuvuuksia - käyttäjälle ymmärrettävällä tavalla [Inmon et al., 1997, s. 183-184; Vassiliadis and Sellis, 1999]. Tunnusomaista OLAP-analysoinnille on tiivistetyn, koko yritystä koskevien tietojen käyttäminen analysoinnin pohjana [Vassiliadis and Sellis, 1999].

Keskikokoisissa ja suurissa yrityksissä käytetään erilaisia ohjelmistotuotteita tapahtumapohjaiseen toimintaan ja analysointiperustaiseen päätöksentekotoimintaan pääosin kahdesta syystä [Thomsen, 1997, s. 9]:

- 1.) Vaativat yritykset tarvitsevat ohjelmiston, joka on mahdollisimman tehokas sekä tapahtumien analysoinnissa, että analysointi-orientoituneessa prosessoinnissa.
- 2.) Nopea päivitys on ehdoton edellytys tapahtumapohjaiseen prosessointiin, kun taas nopea prosessointi ja tiedonhaku ovat ehdottomia edellytyksiä analysointiperustaiseen prosessointiin. Näistä eroista johtuen esimerkiksi tehokas tietokantaperustainen prosessointi vaatii erilaiset lähestymistavat indeksointiin.

Näiden lisäksi erillisiä järjestelmiä voidaan perustella integroidulla näkymällä organisaation tietoon, tiedon muodolla sekä tiedon laadulla [Watson and Gray, 1998, s. 19]. Watson ja Gray [1998, s. 19] toteavat, että tyypillistä käyttäjää ei kiinnosta, mihin tieto on varastoitu. He haluavat päästä siihen käsiksi yhdestä keskitetystä paikasta. Tietovaraston tieto on lisäksi integroitu ja sitä säilytetään standardissa muodossa. He määrittelevät, että tietovaraston tietojen laadukkuuden ansiosta se tarjoaa luotettavan näkemyksen yrityksessä oleviin tietoihin. Käyttäjät haluavat ymmärtää, mitä tieto tarkoittaa, eivätkä selvittää sitä, onko tietty arvo oikea [Watson and Gray, 1998, s. 19]. Toisaalta tapaustutkimuksissani huomasin, että joissakin yrityksissä vain alkuperäisen järjestelmän tieto on käyttäjien mielestä oikeata tai käyttäjät luottavat vain siihen. Tämä ilmenee helposti, jos tietovaraston tieto on yksityiskohtaista tai jos käytössä on ollut aikaisempi järjestelmä tai taulukkolaskentapohjainen ratkaisu tietovaraston korvikkeena.

2.2. Vaikutus ja arvo

Tietovarastoinnin vaikutuksen ja arvon kvantifiointi (määrällinen esittäminen) on vaikeaa. Yleensä vaikutuksia voidaan arvioida vain laadullisen, kvalitatiivisen tutkimuksen perusteella, koska mitattavia asioita tietovarastoinnissa on vähän. Yksi tapa kvantifiointiin olisi kuitenkin päätöksenteon taloudellisten vaikutuksien seuraaminen kahdessa erilaisessa tilanteessa: päätöksenteossa ilman tietovarastoa ja päätöksenteossa tietovaraston kanssa. Toinen mahdollinen kvantifointimahdollisuus olisi erilaisten älykkään liiketoiminnan mittaristojen (engl. scorecard) avulla toteutettava reagointinopeuden mittaaminen, eli esimerkiksi tavaran virheellisyyttä koskevien ilmoitusten eli reklamaatioiden käsittelyyn kulunut aika.

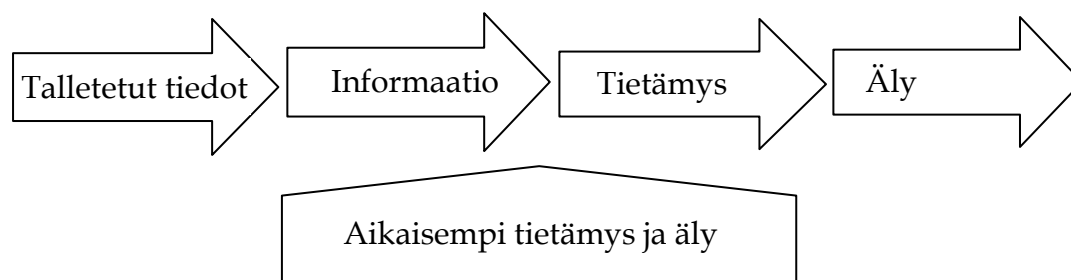
Counihan ja muut [2002] ovat tutkineet tietovarastointihankkeen investoinnin kannattavuutta ja toteavat, että perinteiset arviointimenetelmät ovat rajoittuneita arvioitaessa strategisia sovelluksia kuten tietovarastoa. Counihan ja muut [2002] ovat luoneet investoinnin arviointia varten kehyksen (engl. framework), jolla he ovat pyrkineet arvioimaan investoinnin teknologisia hyötyjä organisaatiovaatimukset huomioiden. Kehyksen ensimmäisessä vaiheessa voidaan korkean tason analysoinnin avulla varmistaa, että tietovaraston toteuttaminen on organisaatiolle välttämätöntä. Korkean tason analysoinnilla he ovat tarkoittaneet taloudellisen ympäristön arviointia ja

toimialan informaatiokeskeisyyttä suhteessa asiakkaisiin. Esimerkkeinä käytännön toimista ovat esimerkiksi tarvittavan tietoteknisen infrastruktuurin laajuuden määrittäminen sekä lyhyt- ja pitkäkestoisten asiakasvaatimusten investointien takaisinmaksuajan arviointi.

Counihanin ja muiden [2002] mukaan toisen vaiheen arviointiprosessi on jaettu neljään vaiheeseen: sitoutumiseen ja tukemiseen, lähestymistavan arviointiin, hyötyjen ajanjaksojen määrittämiseen sekä tekniikoiden arviointiin. He toteavat, että sitoutumisen on lähdettävä jo liiketoimintayksikkö- ja ryhmätasolta. Arvioinnissa on käytettävä joko vaiheittaista, sovelluskohtaista tai koko organisaation laajuista lähestymistapaa, mutta kuitenkin itse arviointi on suoritettava liiketoimintayksikkötasolla. Hyötyjen arviointi on tehtävä lyhyitä ajanjaksoja silmällä pitäen, kuitenkin pitkän tähtäimen arviointia unohtamatta, koska lyhyen tähtäimen hyötyjen määrittäminen on organisaatioille yleensä helpompaa.

Jotta arvioinnista saadaan suurin hyöty, he ovat suositelleet käytettäväksi tapaustutkimukseen perustuvaa lähestymistapaa, jotta investointi tulee arvioitua kokonaisvaltaisesti eikä pelkästään taloudellisista lähtökohdista. He lisäävät, että käytettäessä arvoanalyysiä saadaan käyttäjät osallistumaan arviointiin jo pienessä mittakaavassa ja prototyypitasolla. Lisäksi tarvitaan perusohjeet infrastruktuurin elementtien arviointiin, jotta infrastruktuuriin tehtävät sijoitukset ovat yhdenmukaiset organisaation strategian ja kilpailukyvyyn kanssa. [Counihan *et al.*, 2002]

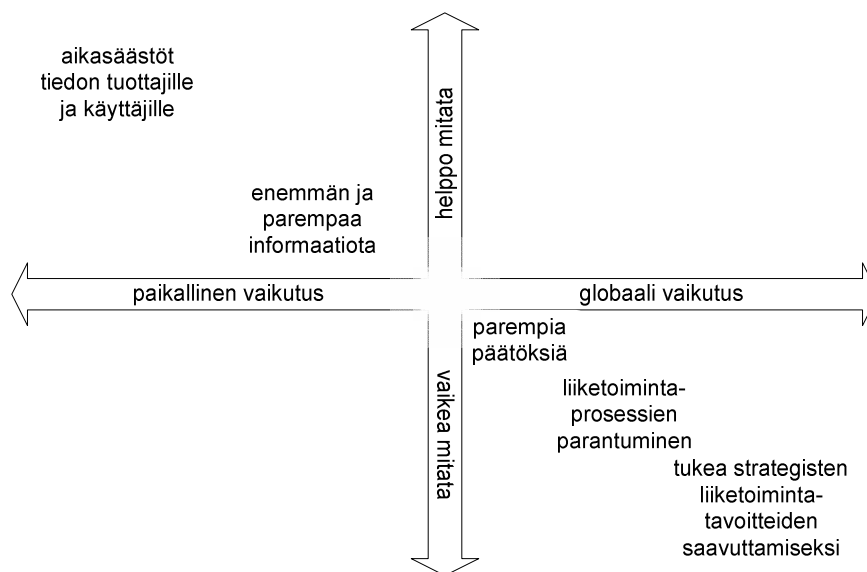
Yrityksen tietovaraston avulla voidaan Gardnerin [1998] mukaan analysoida mitä liiketoiminnalle on tapahtumassa ja syyt, miksi näin tapahtuu. Esimerkiksi yrityksen tulos saattaa pienentyä, koska se valmistaa tuotteita, joiden kysyntä markkinoilla laskee. Tämä myynnin laskeva trendi voidaan esimerkiksi havaita tietovaraston kokoomatiedoista verrattaessa kuluneiden viikkojen myyntiä toisiinsa. Liiketoiminnasta kertyykin paljon tietoa, mutta vain vähän liiketoiminnan suuntaamista avustavaa informaatiota. Gardner [1998] on kuvannut tätä operatiivisten järjestelmien tuottamien tietojen ongelmaa. Operatiivisissa järjestelmissä on talletettuna valtavasti tietoa, mutta tämä talletettu tieto ei kuitenkaan ole Gardnerin mielestä hyödyllistä, ellei sitä pystytä muuntamaan informaatioksi, jolloin talletettuihin tietoihin on liitettynä reaali maailman tulkinta. Talletetuilla tiedoilla pystytään vastaamaan kysymykseen mitä tapahtui, mutta jalostetulla informaatiolla pyritään vastaamaan esimerkiksi kysymykseen, miksi näin tapahtui. Gardner on todennut, että tietovaraston yksi tehtävä on tarjota usean järjestelmän kattavia analysointimahdollisuuksia. Toinen tehtävä on tarjota pääsy yksityiskohtaiseen perustietoon, jolloin vältetään mahdolliset summaamisen aiheuttamat harhat. Tietovarasto on osa tiedonjalostusketjua, joka muuntaa talletettuja tietoja tietämykseksi ja älykkyydeksi. Hannula ja Pirttimäki [2004] ovat kuvassa 2 osuvasti kuvanneet tämän tiedonjalostusketjun liiketoimintatiedon hallinnan näkökulmasta.



Kuva 2. Tiedonjalostusketju liiketoimintatiedon hallinnan näkökulmasta.

Tietovarastoprojektit ovat kuitenkin suuren riskin, mutta samalla suurten tuottomahdollisuuksien hankkeita [Watson *et al.*, 2002]. Watson ja muut [2002] ovat todenneet, että erään tutkimuksen mukaan tietovarastojen keskimääräinen tuottoprosentti kolmen vuoden ajalta oli noin 400 prosenttia koskien yli 60 organisaatiota. He kuitenkin muistuttavat, että tilastoissa ovat yleensä mukana vain onnistuneet projektit. Lisäksi mittaus on suoritettu jälkikäteen. Kuitenkin noin puolet tietovarastohankkeista epäonnistui ensimmäisellä yrityksellä ja epäonnistumisen hinta on kova, koska tutkimuksien mukaan keskimääräinen tietovarastoprojekti maksaa yli miljoona Yhdysvaltain dollaria [Watson *et al.*, 2002]. Toisaalta, jos tietovaraston tietojen perusteella tehdään virheellisiä tai epäonnistuneita liiketoimintapäätöksiä, saattavat seuraukset yritykselle olla vielä ikävämpiä.

Tietovaraston liiketoimintahyödyt voivat olla vaikeasti mitattavia ja hyötyjen vaikutusalue voi vaihdella yksittäisestä käyttäjästä organisaation laajuiseen. Kuvassa 3, Watson ja muut [2002] ovat ryhmitelleet aikaisemmissa tutkimuksissa havaitut hyödyt, vaikutuksen laajuuden ja hyödyn mitattavuuden mukaan neljään kategoriaan. Hyötyjen rahallinen arviointi on kuitenkin erittäin vaikeaa, joten kuvassa 3 esitetyt hyödyt eivät ole yhteismitallisia.



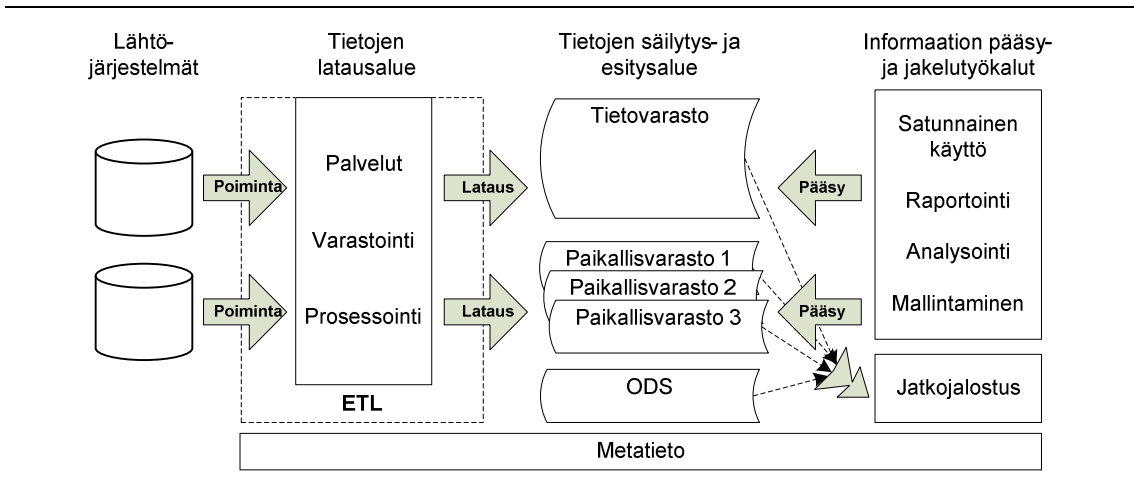
Kuva 3. Watsonin ja muiden [2002] luokittelemat tietovarastoinnin liiketoimintahyödyt kuvattuna kaksiulotteisesti.

Aikahyödyt ja parempi informaatio ovat tietovaraston konkreettisimmat liiketoiminnalliset hyödyt [Watson *et al.*, 2002]. Informaatiota voidaan tulkita paremmin vain, mikäli tietovarasto sisältää tarvittavat tiedot. Watson ja muut [2002] ovat todenneet, että aikahyödyt saavutetaan tiedon nopeamman löytymisen takia. Informaation parantuminen on puolestaan toisen kategorian hyöty, joka voi olla tunnistettavissa. Sen sijaan ei välttämättä ole pääteltävissä, mistä informaation parantuminen johtuu tai mikä on sen mitattavissa oleva hyöty [Watson *et al.*, 2002]. Kolmannessa kategoriassa tunnistetaan tietyt päätökset, jotka ovat selkeästi parempia kuin aikaisemmat päätökset. Watson ja muut [2002] ovat tunnistaneet tämän johtuvan usein tietovaraston parantamista informaation tulkintamahdollisuuksista, koska tietovarasto mahdollistaa tietojen perusteellisemmän analyysin ja johtajat huomaavat tietovaraston avulla muutokset ja trendit nopeammin. Suurimmat hyödyt tietovarastoinnista saadaan, kun tietovarastoja käytetään liiketoimintaprosessien uudelleensuunnitteluun ja liiketoiminnan strategisten tavoitteiden tukemiseen [Watson *et al.*, 2002]. Watson ja muut lisäävät, että nämä hyödyt tarkoittavat yleensä merkittäviä muutoksia organisaation toiminnassa.

2.3. Tietovarastoinnin osa-alueet

Kimball ja Ross [2002, s. 7-15] sekä Ponniah [2001, s. 29] jakavat tietovarastoinnin komponentit neljään osa-alueeseen (kuva 4): *lähtöjärjestelmiin* (engl. source systems), *tietojen latausalueeseen* (engl. data staging area), *tietojen säilytys- ja esitysalueeseen* (engl. data storage and presentation area) ja *tietojen pääsy- ja jakelutyökaluihin* (engl. information access and delivery tools). Tiedot ladataan operatiivisista

lähtöjärjestelmistä, muunnetaan, puhdistetaan, integroidaan ja säilytetään tietovarastossa [Ponniiah, 2001, s. 26].



Kuva 4. Tietovarastoinnin osa-alueet.

Operatiiviset lähtöjärjestelmät keräävät liiketoiminnan tapahtumat (esimerkiksi tietojärjestelmän sisältämät ostolaskut) tietovarastoinnin perustaksi [Kimball and Ross, 2002, s. 7]. Tietovarastoinnilla ei kuitenkaan ole määräysvaltaa tai hallintamahdollisuutta operatiivisten järjestelmien sisältöön tai tiedon muotoon, joten operatiiviset järjestelmät muodostavat Kimballin ja Rossin [2002, s. 7] mielestä tietovarastoinnin ulkoisen osa-alueen. Ponniiah [2001, s. 29-30] on kategorisoinut lähtöjärjestelmät vielä neljään eri tyyppiin, jotka kuvaavat lähteen tietoteknistä luonnetta sekä kerätyn tiedon hankintatapaa. Tuotantotietoa sisältävät yrityksen operatiiviset järjestelmät ovat tyypillisimpiä lähteitä. Sisäiset lähteet ovat esimerkiksi henkilökohtaisia laskutaulukoita, dokumentteja tai osaston paikallisia tietokantoja. Arkistoidut lähteet ovat esimerkiksi operatiivisten järjestelmien varmistusnauhoille tallettamaa historiallista tietoa. Ulkoiset lähteet ovat yleensä erilaisia julkisia yhteenvedoja tai tilastoja, kuten markkinatutkimuksia ja luottotietoja, yrityksen liiketoimintaan liittyvistä asioista, esimerkiksi kilpailijoista.

Tiedon latausvaihe koostuu tiedon keruusta, muuntamisesta ja latauksesta tietovarastoon (engl. extract, transform and load, ETL). ETL on joukko prosesseja, joiden tehtävä on muuntaa operatiivisista järjestelmistä kerätty tieto tietovaraston edellyttämään muotoon. Tiedon latausalue on laajempi suurempi kokonaisuus, kuin pelkkä ETL-prosessointivaihe. Latausalue pitää sisällään lähtöjärjestelmien tiedon varastoinnin siirtotiedostoihin (engl. flat files) ja tietokannan tauluihin. Latausvaiheen tehtävät pitävät yleensä sisällään erilaisia palveluja: tiedon puhdistamista (kuten puuttuvien arvojen käsittely), yhdistämistä (kuten eri järjestelmien asiakkaiden yhdistäminen), dimensioiden ja niiden arvojen yhdenmukaistamista, indeksoimista, laadunvarmistusta sekä julkaisua tietovarastoon. ETL-vaiheen prosessointi tarkoittaa tietojen järjestämistä sekä peräkkäistä prosessointia, jotta esimerkiksi tietojen

keskinäiset riippuvuudet eli tietojen eheys säilyy. [Kimball and Ross, 2002, s. 7-10, 401] Chaudhuri ja Dayal [1997] erottavat edellä mainituista ETL-vaiheen tehtävistä vielä virkistuksen (engl. refresh) eli tietojen päivittämisen tilanteissa, joissa tiedot ovat jo tietovarastossa. Virkistys tapahtuu määräajoin, esimerkiksi päivittäin, jotta lähtöjärjestelmien muutokset saadaan päivitettyä tietovarastoon. Virkistys pitää sisällään myös vanhojen ja tarpeettomien tietojen poiston tietovarastosta.

Käsitelty tieto siirretään erilliseen säilytyspaikkaan, jossa tietoa säilytetään analysointitarpeita varten. Tietoa on yleensä valtavasti, koska historiatietoa saatetaan haluta säilyttää useita vuosia. Säilytettävä tieto on yleensä melko pysyvää, koska analyytikkojen tarvitsemat tiedot ovat tietyltä ajanjaksolta. Tästä syystä tietojen päivityssykli on suunniteltava tarkasti liiketoiminnan ehdoilla. Päivityssykli voi tarpeesta riippuen vaihdella tunnista vuoteen. [Ponniiah, 2001, s. 33-34]

Tietovaraston perustamisen lähestymistavat voidaan jakaa kahteen luokkaan (Taulukko 1): tietovarastoon ja paikallisvarastoon (engl. datamart) [Ponniiah, 2001, s. 25]. Paikallisvarastot ovat loogisia alijoukkoja kokonaisesta tietovarastosta, ja yleensä tiettyyn liiketoiminta-alueeseen kohdennettuja. Tällöin yhdenmukaisella paikallisvarastojen suunnittelulla voidaan välttää niiden pirstoutuminen erilaisten tietojen saarekkeiksi [Ponniiah, 2001, s. 27]. Moody ja Kortink [2000] toteavat, että usein keskitetystä tietovarastosta muodostettujen paikallisvarastojen tiedot ovat suoraan loppukäyttäjien käytettävissä, joten suunnitteluvaatimuksena on tietojen järjestäminen siten, että käyttäjien on helppo ymmärtää ja käyttää niitä. Siksi dimensionaaliset mallintamistekniikat soveltuvat parhaiten paikallisvarastojen mallintamiseen [Moody and Kortink, 2000].

Tietovarasto	Paikallisvarasto
Yhteinen ja yrityksen laajuinen	Liiketoimintayksikkökohtainen
Kaikkien paikallisvarastojen yhdiste	Kuvaa vain yhtä liiketoimintaprosessia
Tiedot saadaan latausalueelta	
Kyselyt yhteisiin esitysalueen tietoihin	Teknologia optimoitu tiedon saantiin ja analysointiin
	Rakenne soveltuu osastokohtaiseen tietojen tarkasteluun
Mallinnustapana voidaan soveltaa perinteistä Chen [1976] esittelemää ER-mallia tai siitä johdettua mallia	Käytössä dimensionaalinen mallintaminen tähtiliitoksien avulla (faktat ja dimensiot). Fakta vastaa mitattavaa asiaa ja dimensio asiahierarkiaa.

Taulukko 1. Tietovaraston ja paikallisvaraston keskeiset erot Ponniiahin [2001, s. 26] näkemyksen mukaan.

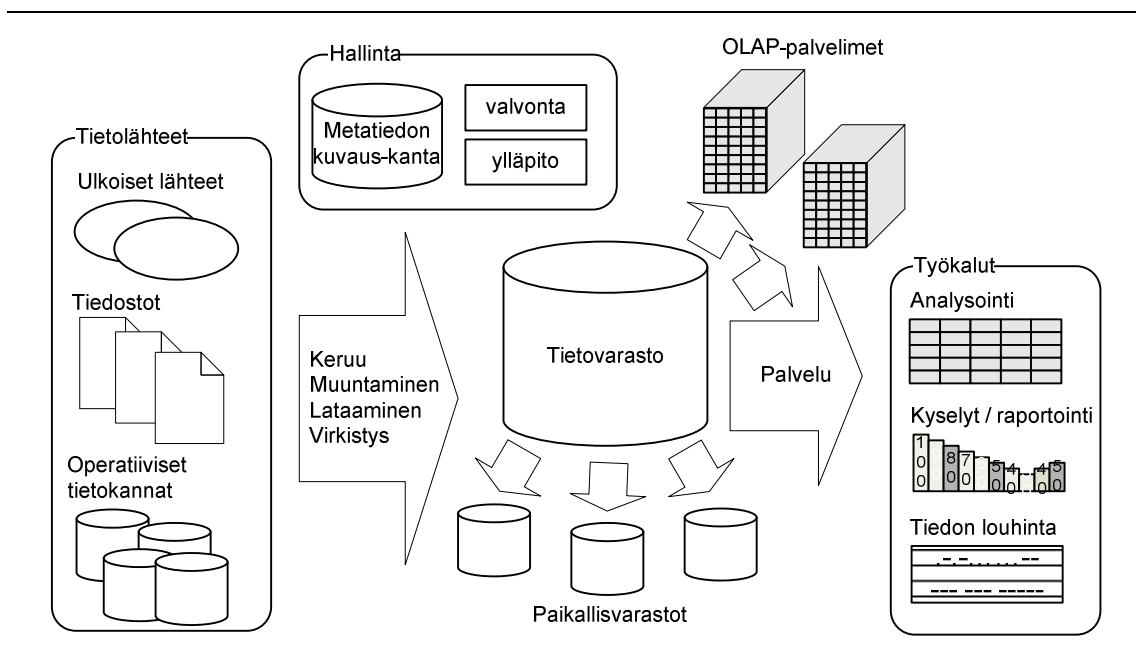
Operatiiviset tietovarastot (engl. operational data store, ODS) ovat eräs tietovaraston muoto, jonka tarkoitus on auttaa toteuttamaan operatiivinen raportointi. Tällaisia tietovarastoja tarvitaan, jos lähtöjärjestelmä ei tue tarvittavaa raportointia. Operatiiviset tietovarastot integroivat ja kopioivat operatiivista tietoa, mutta tiedot päivittyvät harvemmin kuin operatiivisissa järjestelmissä. Operatiivisen tietovaraston raportit ovat yleensä vakioraportteja ja ne vastaavat taktisen päätöksenteon tarpeita, koska tiedot ovat lähes reaaliaikaisia ja melko yksityiskohtaisia. Operatiiviset tietovarastot sijoittuvatkin päätöksenteon näkökulmasta lähtöjärjestelmien ja tietovarastojen väliin. Kimball ja Ross muistuttavat, että operatiivisen tietovaraston hyödyllisyys on organisaatiokohtaista edellä mainituista syistä. [Kimball and Ross, 2002, s. 15-16]

Tietovaraston lisäksi toinen keskeinen komponentti on tiedon esittämiseen ja analysointiin tarvittavat työkalut. Loppukäyttäjän tarpeista ja tietoteknisestä osaamistasosta riippuen työkalut voivat olla esimerkiksi vakioraportteja, ennakoimattomia kyselyjä, moniulotteisen tiedon käsittelyä esimerkiksi tietokuutioiden (engl. data cube, hyper cube) avulla, tilastollista analysointia tai tiedon louhintaa. Vakioraportit ovat yleensä suunnattu aloitteleville käyttäjille ja ennakoimattomat kyselyt edistyneemmille käyttäjille [Ponniah, 2001, s. 34]. Tietoa voidaan myös jatkojalostaa moniin eri tarkoituksiin ja järjestelmiin sopivaksi. Esimerkiksi tiedot siirretään ensin ennustejärjestelmään, josta ne palautuvat takaisin tietovarastoon.

Ponniah [2001, s. 35-36] on todennut, että metatieto ei varsinaisesti ole tietovaraston yksi komponentti vaan se on tietoa tietovaraston sisältämästä tiedosta. Toiset tutkijat pitävät metatietoa kuitenkin kiinteästi tietovaraston komponenttina. Tietovaraston metatieto on tietoa esimerkiksi operatiivisista lähtöjärjestelmistä, ETL-prosessien käsittely- ja liiketoimintasääntöjä sekä loppukäyttäjän tarvitsemää liiketoiminnallista metatietoa [Ponniah, 2001, s. 35-36].

2.4. Tietovarastoarkkitehtuurit

Chaudhuri ja Dayal [1997] esittelevät erään tyypillisen tietovarastointiratkaisun arkkitehtuurin kuvassa 5. Arkkitehtuuri käsittää työkalut datan poimimiseksi operatiivisista ja ulkoisista, yrityksen kontrolloimattomista lähteistä, datan puhdistamiseksi, muuntamiseksi ja integroimiseksi sekä lataamiseksi ja päivittämiseksi tietovarastoon. Tietovaraston arkkitehtuuri voi sisältää myös paikallisvarastoja ja OLAP-kuutioita. Loppukäyttäjille on yleensä tarjolla työkalut tietojen analysointiin, raportointiin ja mahdollisesti tiedon louhintaan. Tietovarastointia tukevat metatietojen kuvauskanta, hallinnan työkalut sekä valvonta- sekä ylläpitotyökalut. [Chaudhuri and Dayal, 1997]



Kuva 5. Eräs tyypillinen tietovarastointiarkkitehtuuri.

3. Tietovaraston tietomallit ja mallintaminen

Organisaation tietomalli on tämän tutkimuksen oleellisin käsite. Oxfordin [2004] tietotekninen sanakirja määrittelee yleisen tietomallin näin:

”Abstrakti malli jostain todellisen maailman tilanteesta tai intressin arvoalueesta, jonka informaatio säilytetään tietokannassa ja jonka looginen malli kuvaa tietokantaan. Termiä tietomalli voidaan myös käyttää joukosta loogisia abstrakteja käsitteitä, joilla muodostetaan sellainen malli.”

Elmasrin ja Navathen mukaan tietomallia käytetään nostamaan käsitetasoa piilottamalla käyttäjiltä teknisiä yksityiskohtia. He kuvaavat tietomallin joukoksi käsitteitä, joita voidaan käyttää kuvaamaan tietokannan rakenne ja sisältö kaaviotasolla. Tietokannan rakenteella he tarkoittavat tietotyyppejä, suhteita ja rajoitteita, joiden pitää vastata tietokannan sisältämää tietoa. [2000, s. 24]

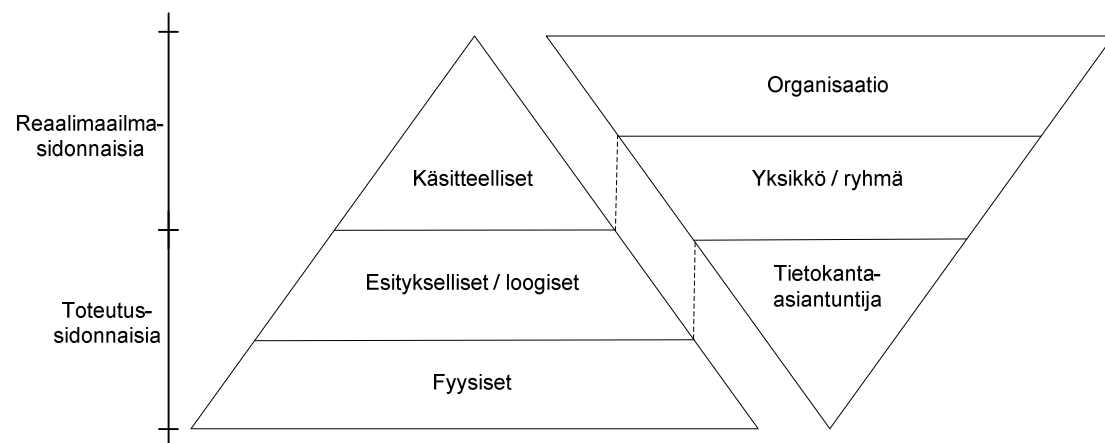
Tietomalleja voidaan kategorisoida monin tavoin. Yksi tapa on jaotella ne sisältämänsä informaation mukaan. Elmasri ja Navathe [2000] jakavat ne tietokannan suunnittelun vaiheina kolmeen tasoon (kuva 6):

- 1.) käsitteellisiin tietomalleihin (engl. conceptual data model, semantic data model),
- 2.) esityksellisiin (engl. logical data model) sekä
- 3.) fyysisiin tietomalleihin (engl. physical data model).

Käsitteelliset mallit tarjoavat tavan esittää käyttäjien näkemykset organisaation tiedoista. Tämä taso on täysin riippumaton alla olevista järjestelmistä. Esitykselliset mallit puolestaan kuvaavat organisaatioon talletetun tiedon tietoyksikköinä, joita käyttäjien on helppo ymmärtää. Esityksellisiin malleihin ei yleensä kuvata fyysisiä

yksityiskohtia, vaan ne määräävät toteutuksessa käytettävän tietokannan loogiset tietoyksiköt ja niiden väliset suhteet. Fyysinen tietomalli taas ei kuvaa tietoa vaan keskittyy tiedon tallentamisperiaatteisiin, kuten B+ -puihin perustuvaan indeksointiin. [Navathe, 1992]

Eritasoiset tietomallit ovat yhteydessä toisiinsa ja niiden avulla voidaan myös kuvata tietomallin sisältämän tiedon laajuutta. Inmon [Inmon *et al.*, 1997, s. 67] yhdistää tietomallien abstraktiotasot organisaation osa-alueisiin (Kuva 6). Käsitteellinen tietomalli liittyy koko organisaatioon, kun taas fyysinen malli saattaa koskea vain yksittäistä asiantuntijaa, joka suunnittelee esimerkiksi toteutuksessa käytettävän tallennusrakenteen. Tässä tutkimuksessa keskityn edellä esitellyn jaottelun ylimpään eli käsitteelliseen tasoon ja miten yrityksen tietomalli vaikuttaa tähän tasoon ja toisin päin, eli miten käsitteellinen mallintaminen liittyy yrityksen tietomalliin.



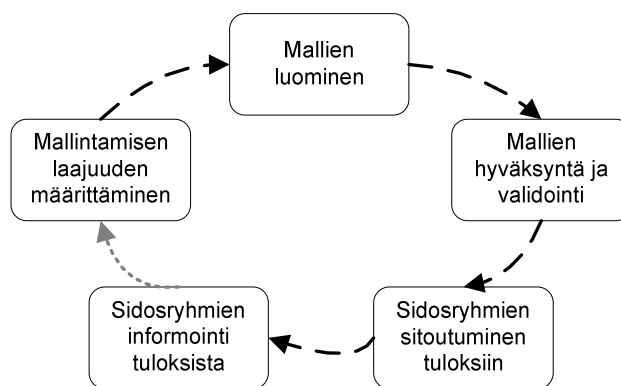
Kuva 6. Laatimani tietomallien jaottelutapa informaation yksityiskohtaisuuden ja tietojen vaikutusalueen perusteella.

Calvanese *et al.* [1999] esittelevät tietovaraston tietojen eräänä mallintamisvaihtoehtona DWQ-lähestymistavan (Data Warehouse Quality). DWQ-malli on samalla tavalla kolmeen tasoon jaettu kuin edellä esitelty luokittelu. Käsitteellinen taso sisältää DWQ-mallissa käsitteellisen esityksen yrityksen tiedoista [Calvanese *et al.*, 1999]. Mallin looginen taso puolestaan sisältää loogisen tietomallin tietolähteiden ja tietovaraston tiedoista. Alin taso DWQ-mallissa on fyysinen taso ja se mallintaa tietojen latauksen tietolähteistä välittäjien (engl. mediator) avulla ja lähteen päälle muodostettavan tietojen kokoajan (engl. wrapper) sekä tietojen säilytykseen liittyvät asiat. Eri tasojen objektit kytketään toisiinsa määrittelemällä niiden välille kuvaukset (engl. mapping) [Calvanese *et al.*, 1999].

Perinteisten tietokantasuunnittelutekniikoiden käyttämisen ongelmana tietovarastointiympäristössä on se, että lopputuloksena on tietokantojen rakenteita, jotka ovat liian monimutkaisia, jotta loppukäyttäjät voisivat niitä ymmärtää ja käyttää [Moody and Kortink, 2000]. Moodin ja Kortinkin [2000] mukaan yksi suuri syy

operatiivisten tietokantojen monimutkaisuuteen on normalisointi, jolla vähennetään tiedon toistoa, mutta tällöin taulujen määrä kasvaa. Tällöin tietojen käyttäjiltä edellytetään navigointikykyä eli esimerkiksi relaatiotietokannoissa semanttisesti yhteenkuuluvien tietojen määrittely on käyttäjän itsensä vastuulla.

Hoppenbrouwers ja muut [2005] ovat tutkineet tietojärjestelmien mallintamistekniikoita tavoitteenaan käsitteellisen mallintamisen uudistaminen kommunikaatio-ohjatuksi tietämyksen muunnosprosessiksi. Mallintamistekniikoiden käytön voi heidän mukaansa jakaa vaiheisiin (Kuva 7), joita ei kuitenkaan tarvitse suorittaa järjestyksessä vaan tarkoituksenmukaisesti soveltaen. Mallintamisen laajuuden määrittäminen on kuitenkin yleensä ensimmäinen vaihe, jolloin määritetään mistä reaali maailman osasta olemme kiinnostuneita.



Kuva 7. Mallintamistekniikoiden käytön vaiheet.

Mallintamistekniikan valinnan jälkeen on määritettävä mallinnettava sovellutusalue (engl. domain) ja sen rajat (engl. constraints). Mallintajan on tiedettävä miksi tiettyä mallintamistapaa käytetään. Oleellista on, että mallintamisen perusteet ja tavoitteet ovat tiedossa etukäteen. Sidosryhmät ovat kiinnostuneita siitä, miten tietty mallintamistekniikka vastaa heidän tarpeisiin. Mallin yksityiskohtaisuuden määrittäminen riippuu mukana olevista sidosryhmistä. Tyypillisesti liiketoimintaihmiset tarvitsevat sovellutusasioiden huomioon ottamista ja sovelluskehittäjät kaipaavat yksityiskohtaisempia ja teknisempiä malleja. [Hoppenbrouwers *et al.*, 2005]

Mallin toteuttamisessa voidaan päivittää myös olemassa olevaa mallia. Osallistujien määrää on rajoitettava ja varsinainen mallintaminen suoritetaan yksilöllisesti. Tällöin työskentelystä saadaan sujuvaa. Sidosryhmäläisten kanssa on suositeltavaa käydä läpi eri vaihtoehtoja ja konkreettisia esimerkkejä yhteisymmärryksen varmistamiseksi. Mallin graafinen esittämistapa on valittava huolella, jotta myös ei-tekniset ihmiset pystyvät tulkitsemaan sitä. Jos malli ei auta uuden tietämyksen hankinnassa, informaation jakamisessa tai johdon tuki lakkaa, on mallintaminen myös osattava keskeyttää. [Hoppenbrouwers *et al.*, 2005]

Malli ja sen edustama esitys tavoitellusta tilanteesta on hyväksyttävä sidosryhmäläisillä, jotta heidät saadaan sitoutumaan sen käyttämiseen. Malli on tällöin

hyväksyttävä kahdesta näkökulmasta: esitystavan puolesta ja lisäksi se on vahvistettava päätöksentekijöiden hyväksynnällä. Hyväksyminen toteutetaan pienryhmissä, jotka koostuvat vain avainhenkilöistä, koska tällöin kerätty palaute on arvokkainta. Hyväksyminen on myös suoritettava kasvotusten, jotta saavutetaan nopeammin yhteisymmärrys mallin sisällöstä ilman, että aikaa kuluu esimerkiksi ylimääräiseen sähköpostilla käytävään täydentävään kommunikointiin. [Hoppenbrouwers *et al.*, 2005]

Kun avainsidosryhmäläiset ovat hyväksyneet mallin, on myös päätöksentekijöiden sitouduttava siihen. Suunnitellun järjestelmän vaikutukset, kulut ja resurssit on hyväksyttävä. Päätöksentekijöille on esitettävä rakennettavan järjestelmän kulut ja hyödyt kompaktisti ja täsmällisesti, jotta he osaavat päättää jatkotoimenpiteistä. Oikeiden ja tärkeimpien sidosryhmäläisten mukaan saaminen on avaintekijä, koska kaikkia lopputulos ei kuitenkaan miellytä johtuen esimerkiksi uuden järjestelmän käyttöönoton vastustamisesta. Tietovaraston mallinnusprosessin lopputulokselle on parempi, jos suunnittelupäätöksiin ja niiden perustaan sitoudutaan mahdollisimman varhain. [Hoppenbrouwers *et al.*, 2005]

Lopuksi muille sidosryhmäläisille on tiedotettava tuloksista, suunnitelmista ja niiden vaikutuksista. Tällöin kulujen ja hyötysuhteen esittely sidosryhmille edistää heidän hyväksynnän saamista uudelle järjestelmälle. [Hoppenbrouwers *et al.*, 2005]

Olennaista koko mallintamisprosessin kannalta on se, että järjestelmän suunnittelussa käytetään liiketoimintaihminen terminologiaa, ei tietojärjestelmäterminologiaa. Mallien lisäksi tarvitaan kuitenkin verbaalia tekstiä järjestelmän kokonaisvaltaisen ymmärryksen varmistamiseksi. Mallit auttavat sidosryhmiä kokemaan, että he ovat osa liiketoimintaa, jota varten järjestelmä rakennetaan. [Hoppenbrouwers *et al.*, 2005]

3.1. Tietovaraston tietomallit

Tietovaraston luominen, kehittäminen ja ylläpitäminen vaativat yhtenäisen käsitteistön, tavan toimia sekä tavan jakaa informaatiota tietovarastoprojektin tekemiseen osallistujien ja sen hyväksikäyttäjien kesken. Näitä tarpeita pyritään tyydyttämään esimerkiksi erilaisten tietomallien avulla. Inmon [Inmon *et al.*, 1997, s. 65] toteaa, että tietovaraston tietomalli koostuu samoista peruskomponenteista kuin operatiivisen järjestelmän käsitelmä. Näitä ovat mm. oliot, suhteet, määritelmät, avaimet, attribuutit, attribuuttien arvoalueet ja ryhmittelyperiaatteet.

Tietovaraston kaksi keskeisintä toteutustapaa perustuvat joko relaatiomalliin (engl. relational model) tai moniulotteiseen malliin (engl. multidimensional model). Moniulotteinen malli voidaan toteuttaa moniulotteisen OLAP-järjestelmän (MOLAP) avulla, joka koostuu moniulotteisesta tietovarastosta ja suorasta OLAP-työkalutuesta. Toinen vaihtoehto on relaatiopohjainen OLAP-järjestelmä (ROLAP), joka koostuu relaatiotietovarastosta (relaatiotietokannasta) ja moniulotteisuutta tukevasta OLAP-moottorista. [Wu and Buchmann, 1997]

Tämän tutkimuksen tekninen ideologinen tausta on relationaalinen, vaikka tutkimuksessani pyrinkin sivuuttamaan tekniset toteutukset. Koska tutkimukseni ei ota kantaa toteutustapoihin, voidaan tutkimuksen tuloksia soveltaa kumpaankin toteutusteknologiaan.

Wu ja Buchmann [1997] huomauttavat samasta asiasta, johon itsekin tutkimusta tehdessäni törmäsin, eli siihen, että tietovaraston mallintaminen on monimutkainen asia ja siinä syntyy helposti vääriä käsityksiä. Tutkijat eivät myöskään ole päässeet yksimielisyyteen tietovaraston mallintamismenetelmästä eivätkä edes käytettävästä lähestymistavasta. He toteavat, että monia väärinkäsityksiä on syntynyt, koska käsitteellistä ja loogista suunnittelua ei ole erotettu toisistaan. He huomauttavat, että moniulotteisen mallin käyttäminen käsitteelliseen mallintamiseen ei tarkoita, että tietovarasto myös toteutettaisiin moniulotteisena tietokantana, vaan se voidaan toteuttaa yhtä hyvin relaatiotietokantana.

Operatiivisista järjestelmissä oleva tietomassa on suodatettava ja yhdenmukaistettava tietovarastoon soveltuvaksi. Inmon ja muut [Inmon *et al.*, 1997, s. 66] toteavat, että yritystietomallin on täytettävä laaja-alaiset informaatiotarpeet, jolloin esimerkiksi tietojen käytettävyyssarvo kasvaa. Yleensä tämä laajempi hyödyntämistarve syntyy koko yhtiön tai sen tietyn osan tarpeista. Tästä on seurauksena se, että yritystietomalli on muodostettava korkealla abstraktiotasolla. Inmon esittelee joitakin keskeisiä toimenpiteitä muodostettaessa yritystietomallista tietovarastomalli:

- Yritystietomallista poistetaan operatiivinen tieto.
- Tiedot liitetään aikaan.
- Talletetuista tiedoista johdetut tiedot lisätään tietovaraston tietomalliin.
- Tietovaraston tietomallissa kuvatuille käsitteille suoritetaan pysyvyyssanalyysi, jolla selvitetään tietojen säilytysaika tietovarastossa.
- Tietojen suhteet konkretisoidaan esimerkiksi hierarkisoimalla ulottuvuudet.

Yritystietomallin operatiivisten tietojen valinnalla tarkoitetaan tietomallista valittavia ominaisuuksia, toisin sanoen poistetaan ne käsitteet, joita ei tarvita tietovarastossa. Myös valittavien tietojen tietosisällöstä suodatetaan vain tarpeellinen informaatio. Esimerkiksi lisäkuvauskentät, joilla suoritetaan järjestelmän parametrisointia, ovat tietovaraston kannalta tarpeettomia. Lopputuloksena pitäisi olla vain sellaista tietoa, jota loppukäyttäjä hyväksikäyttää tai tietoa, josta johdetaan päätöksentekoa avustavaa tietoa.

Historiallisen perspektiivin saavuttamiseksi jokaiseen tietovaraston tietueeseen on liitettävä aikatieto, koska tietovaraston tiedot ovat aina aikaan sidottuja. Aikatieto voidaan tallentaa aikaleimana tai tieto voidaan esimerkiksi summata kuukausitasolle. Aikatiedot mallinnetaan aina dimension hierarkiakuvaukseen. Dimensio eli ulottuvuus

on keskeinen käsite liiketoiminnassa (esimerkiksi aika tai asiakas), johon perustuen voidaan muodostaa yhteenvedotietoja.

Jotta tietovarastosta saadaan selkeä ja ymmärrettävä koko organisaatiolle, tiedot yhdenmukaistetaan ja muutetaan yhteismitallisiksi, esimerkiksi valuuttojen yhteismitallistaminen. Samoin tuotteen pituuden muuttaminen tuumista senttimetreiksi tai tuotteen pituuden muuttaminen metreistä senttimetreiksi on välttämätöntä. Alkuperäisestä tiedosta ja merkityksestä voidaan myös johtaa uusia merkityksiä sekä uutta tietoa. Esimerkiksi tuotteen koodin perusteella voidaan päätellä tuotteen valmistuspaikka (ilmaisee onko tuote omalla tehtaalla valmistettava vai ei), toisin sanoen kyseessä on johdettu tieto.

Tietomallin pysyvyysanalyysin lopputuloksena voi olla esimerkiksi vuosi vai kymmenen vuotta. Tällä tiedolla on sitä suurempi merkitys, mitä nopeammin ja enemmän tietovarastoon kertyy tietoa. Tietovaraston tietomassat voivat kasvaa erittäin suuriksi. Toisaalta on oleellista selvittää, mikä on tiedon hyödyntämistarve myöhemmin, jotta tietoa ei säilytetä liian vähän aikaa.

Lisäksi eri tietojen keskinäiset suhteet on mallinnettava. Tiedot liitetään toisiinsa siten, että suhde heijastaa niiden suhdetta reaali maailmassa sekä määrittellään olioiden väliset keskinäiset esiintymissuhteet (esimerkiksi yhden suhde moneen).

3.2. Käsitteellinen mallintaminen

Mallintamisella tarkoitetaan käsitetehtävää, joka johtaa kuvaukseen, jollain tietyllä medialla, joko fyysisenä ilmentymänä tai mallinnetun käsitteen esityksenä [Martin *et al.*, 2004]. Käsitteellistä mallintamista voidaan nimittää myös *semanttiseksi* mallintamiseksi. Tietovaraston käsitteellinen malli on Calvanesen ja muiden [Calvanese *et al.*, 1999] näkemyksen mukaan abstrakti esitys yrityksen kontrolloimista tiedoista ja jota on täydennetty lähtötietojen käsitteellisellä kuvauksella sekä tietovaraston tietoja kuvaavien käsitteiden ja niiden välisten suhteiden määrittämisellä. Käsitteellistä mallia voidaan kuvata esimerkiksi laajennetun ER-mallin (Entity-Relationship-model) avulla [Calvanese *et al.*, 1999]. ER-malli koostuu Calvanesen näkemyksen mukaan moniulotteisista olioista, joilla on monia ominaisuuksia, ja suhteista, joiden keskinäiset riippuvuudet voidaan kuvata mallin avulla. Lisäksi Calvanesen määritelmässä käsitteellinen kuvaus sisältää sovellusalueiden (engl. domain) oliot, niiden attribuutit ja olioiden väliset suhteet.

Martin ja muut [2004], kokoavat mallintamisen keskeiset pyrkimykset kuuteen perustavoitteeseen, joiden avulla voidaan eri lähestymistapoja arvioida ja verrata:

Mallintamisen tarkoituksena on ihmisten välinen kommunikointi. Mallit ovat formaaleja artefakteja eli ne ovat ihmisten kehittämiä ja käyttämiä. Tämän takia mallintamisen pitää avustaa ihmisten välistä vuorovaikutusta.

Mallintaminen on monimutkaisuuden kompromissi. Mallintaminen on todellisuuden pelkistämistä ja sen tarkoitus on monimutkaisuuden hallinta.

Nimeäminen mallintamisessa on merkityksellistä. Käsitteitä ja niiden välisiä suhteita nimeämällä yhdistetään muodolliset artefaktit ja niihin liitetään semanttiset tulkinnat.

Mallintamisella on usein liikaa metatasoja. Meta on suhteellinen termi, ei absoluuttinen, joten mallien olioiden ilmentymät ovat kytkettävä myös reaali maailman asioihin. Mallin yleisillä primitiiveillä esitetään tietyt reaali maailman oliot ja niiden väliset yhteydet.

Mallin asioiden riippuvuus ei ole aikasidonnaista. Jos asiasta seuraa toinen asia, nämä asiat eivät välttämättä seuraa toisiaan myös ajallisesti. Esimerkiksi myyntitilauksesta seuraa yleensä tuotteen toimitus, jota seuraa laskutus. Laskutus voi kuitenkin tapahtua jo ennen tuotteen toimitustakin.

Mallintamisen metodologia ei saa häivyttää tietojen välistä arkkitehtuuria. Arkkitehtuuri pitää ilmetä, vaikka metodit ja teknologiat muuttuisivat.

Tietovarastojen käsitteellistä mallintamista on tutkittu pääasiassa kahdesta näkökulmasta: dimensionaalisen mallintamisen ja ETL-prosessien mallintamisen kautta [Rizzi *et al.*, 2006]. Vaikka useita käsitteellisiä malleja on esitetty, ei yhtäkään niistä ole toistaiseksi hyväksytty standardiksi tietovaraston mallintamiseen ja lisäksi kaupallisilla toimittajilla on omat suunnittelutapansa [Rizzi *et al.*, 2006].

3.2.1. ER-mallintaminen

Moody ja Kortink [2000] väittävät, että perinteiset mallintamistekniikat, kuten ER-mallintaminen, soveltuvat parhaiten keskitetyn tietovaraston suunnitteluun, jolloin relaatiotietokannan normalisoitu tietokantasuunnittelu varmistaa datan yhtenäisyyden ja eheyden. He korostavat, että esimerkiksi tietojen raju summaaminen hukkaa tässä vaiheessa informaatiota ja rajoittaa analysointimahdollisuuksia. He toteavat, että yrityksen tietomallin pitää olla keskitetyn tietovaraston perusta mallintamistavasta riippumatta.

Keskityn tässä tutkimuksessa kuitenkin dimensionaaliseen mallintamiseen, joka soveltuu mielestäni paremmin yhteistyössä liiketoimintaihmissa kanssa tehtävään tietovaraston tietojen mallintamiseen, koska käsitteiden suhteiden keskinäinen riippuvuus tulee paremmin tällöin esille.

3.2.2. Dimensionaalinen mallintaminen

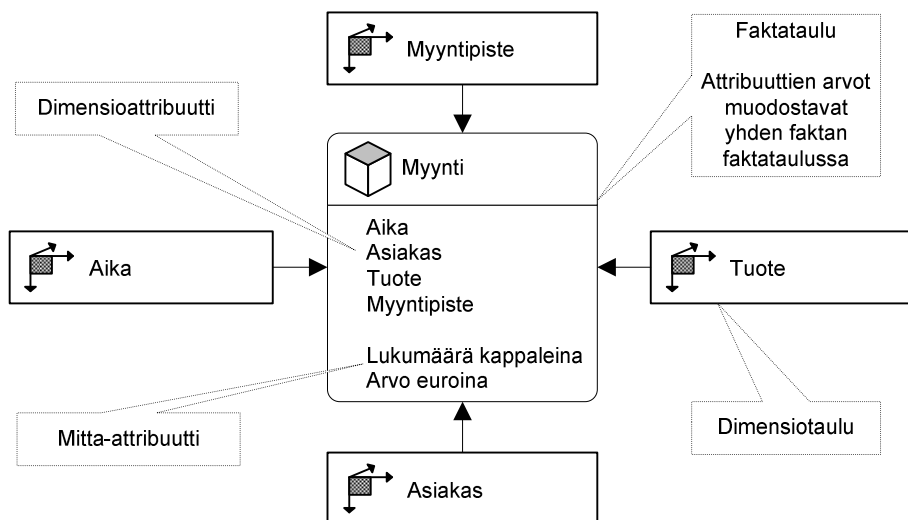
Dimensionaalisen mallintamisen tavoite on tuottaa tietokantarakenteita, jotka loppukäyttäjien on helppo tulkita ja käyttää tietokantakyselyiden muodostamiseen. Toinen tavoite on maksimoida kyselyiden muodostamisen ja suorittamisen tehokkuus.

Tehokkuuden parantumiseen pyritään taulujen ja tarvittavien liitosten määrän vähentämisellä. [Moody and Kortink, 2000]

Dimensionaalinen mallintaminen koostuu faktoista ja dimensioista. Faktaa käytetään tallentamaan mittoja tai tiloja koskien tapahtumaa tai tilaa, missä tietyn mitta-attribuutin arvo riippuu dimensioattribuuttien arvoista [Schneider, 2003]. Tällä tavalla organisoitu tieto on esimerkiksi myytyjen tuotteiden (dimensioattribuutti) määrä (mitta-attribuutti) [Moody and Kortink, 2000]. Mittoja ja tiloja voidaan analysoida erilaisten kriteerien perusteella, jotka ovat organisoitu dimensioihin [Schneider, 2003]. Fakta kuvataan joukkona attribuutteja, joista osa on mitta-attribuutteja ja osa dimensioattribuutteja [Schneider, 2003]. Lisäksi faktalla voi olla attribuutteja, joilla viitataan tiettyyn dimensioon liittyvään informaatioon (dimensiotaulu) [Schneider, 2003]. Faktataulu on linkitetty dimensiotauluihin yhden suhde moneen -suhteella [Moody and Kortink, 2000]. Faktatauluissa ilmaistaan pelkistetysti minkä dimensioattribuuttien mukaan mitta-attribuutit ovat organisoitu. Edellä mainitut dimensionaalisen mallintamisen keskeisimmät käsitteet on kuvattuna kuvassa 8.

Dimensiot puolestaan ovat hierarkkisia rakenteita, jotka koostuvat useasta asiasta, käsitteestä tai oliosta, joilla on kiinteä keskinäinen koostesuhde. Tällainen on esimerkiksi aikahierarkia, joka koostuu päivästä, kuukaudesta ja vuodesta.

Dimensionaalisen mallintamisen perusrakenneosa on *tähtimalli*. Tähtimalli koostuu yhdestä laajasta tietokannan faktataulusta ja useista dimensiotauluista. Fakta muodostaa tähden keskustan ja dimensiot muodostavat tähden sakarat (kuva 8). Tähtimallin voi toteuttaa, joko erityisillä OLAP-työkaluilla tai perinteisillä relaatiotietokantajärjestelmillä. [Moody and Kortink, 2000]

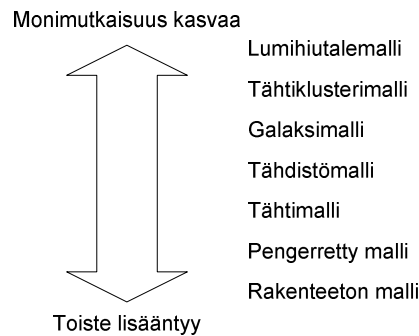


Kuva 8. Tähtimalli kuvaa faktan ja dimensiot tähden muotoisella rakenteella.

Tähtimallin dimensiotaulut eivät yleensä ole normalisoituja ja ne koostuvat yhdestä tai useammasta sisäänrakennetusta hierarkiasta [Moody and Kortink, 2000]. Esimerkiksi dimensio, asiakas, voi sisältää kuvauksen teollisuushierarkiasta ja

sijaintihierarkiasta kertomalla mihin teollisuuden haaraan ja mihin maantieteelliseen alueeseen asiakas sijoittuu.

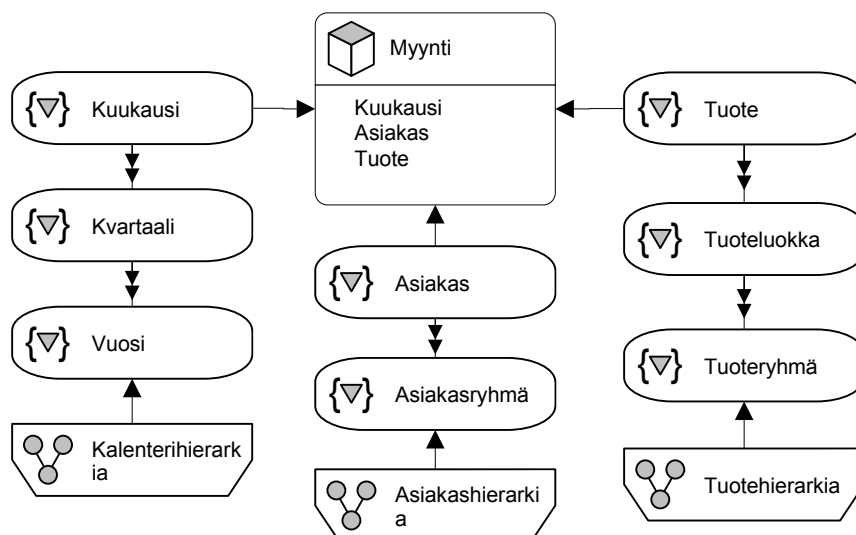
Dimensionaalisten mallien muodostamiseen ER-mallista on monia mahdollisuuksia. Moody ja Kortink [2000] esittelevät joidenkin vaihtoehtojen suhdetta tarvittavien taulujen määrään, josta seuraa monimutkaisuutta tai tietojen toistumista (kuva 9).



Kuva 9. Dimensionaalisten suunnittelumallien vaikutus monimutkaisuuteen ja toisteeseen.

Rakenteeton malli (engl. flat schema) on yksinkertaisin mahdollinen malli, kadottamatta informaatiota. Tällöin jokaiseen tietomallin aiheeseen liitetään kaikki siihen liittyvät muut aiheet. Tämä malli voidaan toteuttaa rakenteettomana tiedostona. Pengerrettyssä mallissa (engl. terraced schema) jokaista tapahtumaa vastaa yksi taulu ja hierarkioina ovat vain tapahtumaan liittyvät välittömät käsitteet. Tähdistömalli (engl. constellation schema) koostuu joukosta tähtimalleja, jotka ovat hierarkkisesti linkitetty faktatauluihin. Galaksimalli (engl. galaxy schema) on joukko tähtimalleja tai tähdistömallia, joilla on jaetut dimensiot, mutta faktataulujen ei tarvitse kuulua suoraan yhteen. [Moody and Kortink, 2000]

Lumihiutalemallissa (engl. snowflake schema) hierarkiat eivät ole normalisoituja, joka on yksi esimerkki siitä, että dimensionaalisisessa mallintamisessa käsitteellisen ja loogisen mallintamisen rajat ovat häilyviä (esimerkiksi verrattaessa tähti ja lumihiutalemallia).



Kuva 10. Esimerkki dimensionaalisesta tähtimallista auki purettuna.

Moody ja Kortink [2000] esittelevät metodin, jolla dimensionaaliset mallit muodostetaan perinteisestä ER-mallista. Menetelmän ensimmäisessä vaiheessa tietomallin kohteet luokitellaan kategorioihin: tapahtuma-, komponentti- ja luokitteluolioihin. Tapahtuma-oliot ovat liiketoiminnan tapahtumia kuten tilauksia. Komponentti-oliot liittyvät tapahtumiin vastaamalla esimerkiksi kysymyksiin kuka, mitä ja milloin. Esimerkiksi asiakas ja tuote ovat tällaisia olioita. Luokitteluoliot ovat tyyppitietoa, esimerkiksi asiakkaan tyyppi tai ryhmä. Toisessa vaiheessa tunnistetaan mallin dimensiohierarkiat, jolloin tunnistetaan olioiden väliset riippuvuudet ja suhteen laatu hierarkkisesti esimerkiksi lapsi-vanhempi-suhteen tavoin. Kolmannessa vaiheessa hierarkiat supistetaan, jolloin esimerkiksi paikka sisältää maakunnan ja kunnan. Samassa vaiheessa tapahtumien numeerisista mitta-arvoista valitaan yhteen laskettavat attribuutit ryhmiteltävien tietojen suhteen. Esimerkiksi myydyistä tuotteista muodostetaan uusi olio, joka sisältää kokonaisymyynnin rahallisen arvon ja keskimääräisen myynnin. Lopputuloksesta muodostuu dimensionaalinen malli [Moody and Kortink, 2000].

Keskitetystä tietovarastosta muodostettujen paikallisvarastojen tiedot ovat suoraan loppukäyttäjien käytettävissä, joten suunnitteluvaatimuksena on tietojen järjestäminen siten, että käyttäjien on helppo ymmärtää ja käyttää niitä. Siksi dimensionaaliset mallintamistekniikat soveltuvat parhaiten paikallisvarastojen mallintamiseen. [Moody and Kortink, 2000]

3.2.3. Muita malleja

Läheisesti käsitteelliseen mallintamiseen ja siihen perustuvaan yrityksen tietomalliin liittyy keskeisten liiketoimintatermien avulla mallinnettu käsittemalli (engl. subject model). Käsittemalli on eräänlainen käsittekaavio, joka kuvaa keskeiset käsitteet ja niiden

väliset suhteet. Käsitelmä ei ole vakiintunut vaihe tietovarastointiprosessissa vaikka useat lähteet sitä eri tavoin hyödyntävätkin kuten [Kendle, 2005; Ballard et al., 1998].

Kendle [2005] on esitellyt käsitelmän (engl. enterprise subject area model), jonka avulla organisaation keskeiset käsitteet voidaan tunnistaa ja luokitella. Ensin kerätään koko organisaation keskeiset korkean tason käsitteet ja etsitään toisistaan riippuvat käsitteet. Esimerkkikäsitteitä voisivat olla asiakas ja varaus, joiden välillä on keskinäinen riippuvuus. Inmon [2002, s. 361] on suositellut, että käsitteanalyysissä valitaan dominoivaksi käsitteeksi riittävän suuri käsite, jotta se on merkityksellinen, mutta riittävän pieni, jotta se on helposti toteutettavissa. Seuraavaksi Kendle on luokitellut löydetty käsitteet kolmeen ryhmään liiketoiminnan näkökulmasta: liikevaihtoon, toimintoihin ja tukeen. Liikevaihtoa tuottava käsite on esimerkiksi myynti. Toimintona on esimerkiksi lento ja tukikäsitteenä asiakas. Lisäksi Kendle on jakanut käsitteiden sisältämän tiedon elämänsäkaaren perusteella perus-, tapahtuma- ja informaatioluokkiin. Peruskäsitteet, kuten toimipaikka, määrittävät ja tukevat muuta tietoa. Tapahtumakäsitteiden tiedot syntyvät päivittäisen liiketoiminnan seurauksena. Tällainen käsite on esimerkiksi lennon varaus. Informatiivisen käsitteen tieto on historia-, summa- ja perittyä tietoa operatiivisista tiedoista. Tähän ryhmään kuuluvat esimerkiksi hinnoittelu ja aikataulu. Mallin lopputuloksena syntyy graafinen kaavio ja metatiedon kuvauskannan keskeisin sisältö.

3.3. Looginen mallintaminen

Looginen mallintaminen on liiketoiminnan analysointitekniikka. Jokaista yrityksen liiketoimintoa käsitellään tietyllä tavalla, esimerkiksi myyntiorganisaatio käsittelee tuotteiden myyntitietoa selvittääkseen, kuinka paljon tuotteita on myyty viime kuussa ja ennustavat tulevan kuukauden myyntimäärää. Looginen tietomalli dokumentoi loogisten tietojen suhteet riippumatta siitä, miten ne ovat toteutettu fyysisissä tietokannoissa ja ohjelmissa. [Moss and Atre, 2003, s. 73]

Yhtenä loogisen tietomallin tarkoituksena Adelman ja Moss [2000, s. 231] pitävät organisaation liiketoiminnan dokumentoimista tietoarkkitehtuuriksi, esimerkiksi tietovaraston rakentamista varten. He eivät kuitenkaan tee selkeätä eroa loogisen ja semanttisen tietomallin välillä, joten mallintamisen vaiheiden määrittäminen on hankalaa. Tällainen jaottelu on kirjallisuudessa valitettavan yleistä.

Eräs looginen tietomalli on nimeltään relaatiomalli, jossa tietojen käsittelemiseen ja esittämiseen sovelletaan joukko-oppiin perustuvaa relaatiotietokantateoriaa. Relaatiomalli kuuluu esitykselliseen tasoon. Kimball ja Ross [2002] esittelevät moniulotteisen mallinnuksen, joka on myös yksi tapa mallintaa tietovarasto.

Thomsen [1997] vertaa relaatiomallin ja moniulotteisen mallin rakennetta. Loogiseen tietokantasuunnitteluun käytetty relaatiomalli koostuu attribuuttien arvojoukoista, attribuuteista, monikoista, relaatioista ja kaavioista. Moniulotteinen malli esitetään ulottuvuuksilla, hierarkioilla, kuutioilla ja kaavioilla. Relaatiomalli keskittyy

perinteisiin relationaalsiin operaatioihin, kuten liitokseen ja karteesisen tuloon. Moniulotteisessa mallissa pääpaino on tieto-operaatioissa, kuten tietokuutioiden tietojen käsittely kaavojen avulla. [Thomsen, 1997, s. 237, 464]. Lisäksi tietokuutioiden tietojen jalostamiseen on oma operaatiojoukko kuten viipalointi (engl. slice and dice) ja karkeistava analyysi (engl. roll-up) [Chaudhuri and Dayal, 1997]. Viipaloinnilla tietoja voidaan dimensioiden osajoukon suhteen yhdistellä ja erotella (projektio). Karkeistavalla analyysillä tietoja yleistetään suuremmiksi kokonaisuuksiksi liikkumalla dimensiohierarkiassa ylöspäin [Chaudhuri and Dayal, 1997].

Moniulotteinen mallintaminen on melko uusi asia, koska vielä vuonna 1997 Thomsen [1997, s. 463] toteaa, että moniulotteista formaalia mallia ei ole olemassa. Thomsen esittelee kuitenkin joitakin moniulotteisten mallien ominaispiirteitä, jotka moniulotteisen mallin pitää täyttää. Mallin on mahdollistettava monitasoiset ulottuvuudet. Mallin on tuettava kaavoja (engl. formulas), joilla ulottuvuuksien jäsenet liitetään ulottuvuuksien sisäisiin ja välisiin toimintoihin. [Thomsen, 1997, s. 450]

Tietovarastoa käytetään analyyttiseen prosessointiin, mikä vaatii moniulotteisuutta. Siksi tietovaraston mallintamiseen käytetyn mallin on kyettävä ja sovittava moniulotteisuuden kuvaamiseen. Fakta ja dimensio ovat moniulotteisen mallin perusrakenteita. Wu ja Buchmann [1997] toteavat, että moniulotteinen malli tarjoaa selkeän ja luonnollisen näkymän tietoihin, joten mallia voi käyttää myös käsitteelliseen mallintamiseen. [Wu and Buchmann, 1997]

Eri mallien rakenteilla voi olla analogioita keskenään, vaikka ne olisivat täysin erityyppisiä. Wu ja Buchmann [1997] ovat linkittäneet semanttisen ER-tietomallin, moniulotteisen, ja loogisen relaatiomallin rakenteet toisiinsa taulukossa 2 kuvatulla tavalla. Vaikka linkitys mallien välillä onkin mahdollista, mallit eivät ole yhtä soveltuvia eri tilanteisiin. Esimerkiksi ER-malli ei ole hyvä analyyttiseen prosessoinnin mallintamiseen, koska malli korostaa, että ensin tunnistetaan käsitteet ja sen jälkeen määritetään niiden keskinäiset suhteet ja saman periaatteen mukaan moniulotteisen mallin dimensiot ja faktat. Kuitenkin yleensä analyttisessä mallintamisessa ensin määritetään analysoinnin oliot ja sen jälkeen niihin liittyvä ulottuvuudet. Relaatiomallin formaalien määritelmien perusteella malli tukee melko hyvin moniulotteisuutta. Se ei kuitenkaan sovellu käsitteelliseen mallintamiseen, koska analyttikot yleensä näkevät tiedot eli mitta-attribuutit analysoitavina kohteina niihin liittyvine dimensioineen. Relaatiomalli on kuitenkin hyvä ehdokas tietovaraston loogiseen suunnitteluun luontaisen moniulotteisuuden ja kyselyiden prosessoinnin matemaattisen pohjan perusteella. [Wu and Buchmann, 1997]

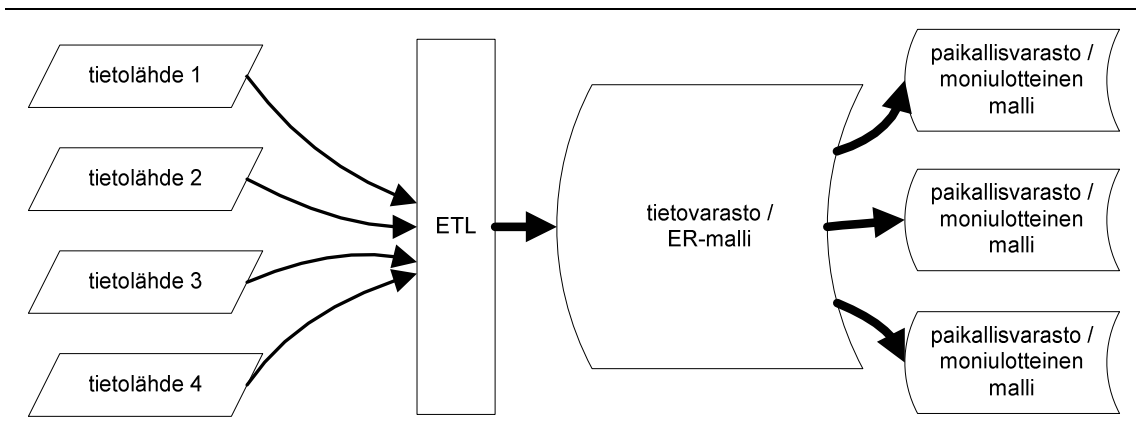
ER malli	Moniulotteinen malli	Relaatiomalli
suhde	fakta	relaatio
käsite	ulottuvuus	

Taulukko 2. ER-mallin, moniulotteisen ja relaatiomallin rakenteiden vastaavuudet [Wu and Buchmann, 1997]

Wu ja Buchmann [1997] suosittavat, että tietovaraston mallintamiseen käytetään yleisesti tietojen mallintamisessa käytettävää kolmikerroksista mallintamistapaa, käsitteellistä, loogista ja fyysistä. He ehdottavat, että moniulotteista mallia käytetään käsitteelliseen mallintamiseen ja relaatiomallia loogiseen mallintamiseen.

3.4. Mallintamisen vaikutuksia tietovaraston arkkitehtuuriin

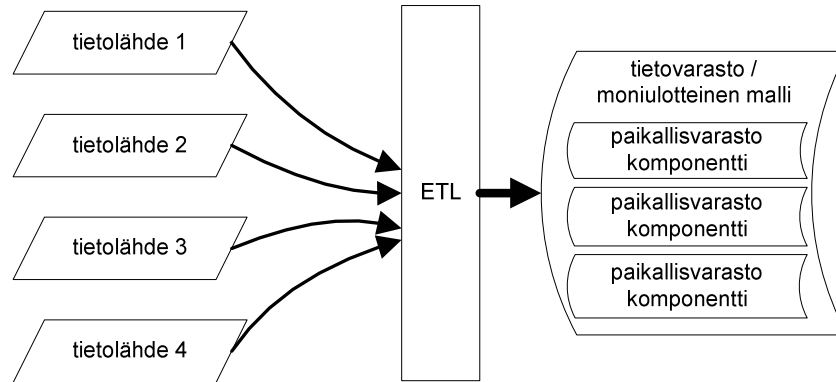
Jukic [2006] esittelee kolme tietovarastojen ja paikallisvarastojen (engl. data mart) mallintamistapaa. Jukic toteaa, että Inmon [2002] esitteli ensimmäisenä tietovaraston integroituna tietokantana, joka on mallinnettu perinteisellä reaali maailman mallintamistekniikalla, ER-mallilla. Kun tietovarasto on luotu, se toimii lähtökohtana moniulotteisesti mallinnetuille paikallisvarastoille sekä muille analyttisille järjestelmille, kuten kuvasta 11 ilmenee. Keskitetty tietovarasto toimii integroinnin keskuksena. Kun tietovaraston malli on kunnossa, on ETL-prosessin vaatimukset selvitetty ja sen toteuttaminen on suoraviivaista. Tietovarastossa olevia tietoja voidaan myös hyödyntää esimerkiksi tiedonlouhinnassa. Keskitetystä lähestymistavasta käytetään myös laajemman kokonaisuuden kattavaa nimeä yrityksen informaatiotehdas (engl. Corporate Information Factory, CIF). [Jukic, 2006]



Kuva 11. ER-mallinnettu tietovarasto [Jukic, 2006].

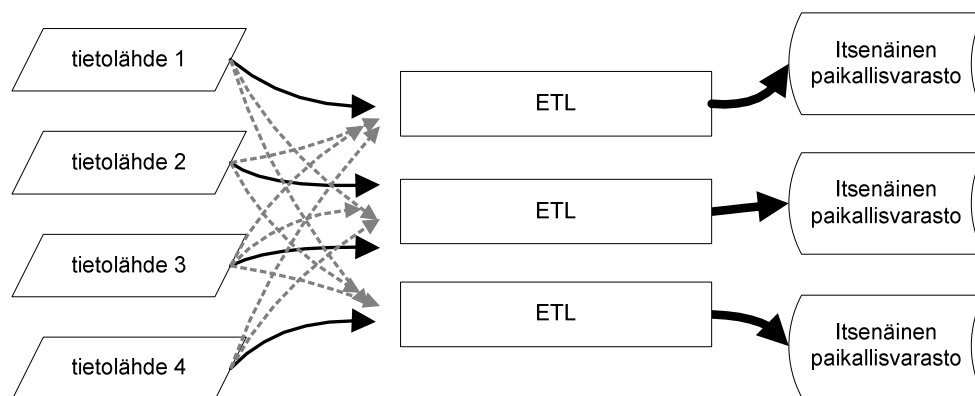
Toinen lähestymistapa on Kimballin ja Rossin [2002] esittämä, joissa tietovarasto on kokoelma moniulotteisesti mallinnettuja paikallisvarastoja (kuva 12). Lähestymistapa on tietojen keräyksen osalta samanlainen kuin edellisessä kappaleessa kuvatussa Inmonin mallissa, mutta tietovaraston mallintamistekniikka on erilainen. Tässä mallissa paikallisvarastot sisältyvät tietovarastoon, kun edellisessä mallissa ne

johdettiin tietovarastosta. Käytettäessä tätä lähestymistapaa, aluksi suunnitellaan joukko yleisesti käytettyjä ulottuvuuksia, yhdenmukaiset ulottuvuudet (engl. conformed dimensions), jotka ovat yhteiset koko tietovarastolle. Tämän jälkeen malliin lisätään oliot, reaali maailman käsitteitä vastaavat faktat. Mallissa jokainen faktataulu on kytketty useisiin ulottuvuuksiin ja joitakin ulottuvuuksia on jaettu usean faktataulun kesken eli esimerkiksi tuote- ja aikadimensiot liittyvät myyntiin. [Jukic, 2006]



Kuva 12. Moniulotteisesti mallinnettu tietovarasto [Jukic, 2006].

Kolmannessa lähestymistavassa (kuva 13) itsenäiset paikallisvarastot luodaan riippumattomasti muista organisaation paikallisvarastoista. Tätä suunnittelutapaa ei kuitenkaan suositella, koska ETL-mallintamiseen tulee turhaa toistoa ja organisaationkattava analysointi sekä kommunikointi vaikeutuvat. Ongelmia tulee myös, mikäli tulevaisuudessa on tarvetta integroida itsenäisissä tietovarastoissa olevia tietoja keskenään. Huolimatta lähestymistavan huonoista puolista, itsenäinen lähestymistapa on käytössä useiden yritysten tietovarastointiratkaisuissa. Syynä näyttäisi olevan yrityksen laajuisen näkökulman puuttuminen tietojen analysointivaiheessa. Tällöin kokonaisnäkemysten saavuttamiseksi on kerättävä yhteen yrityksen kokonaiskustannukset eli liiketoimintasektoreiden kustannukset. Organisaatioiden eri osastot ovat luoneet omia analysointitarpeita silmällä pitäen omat paikallisvarastot, joissa on tietoa esimerkiksi henkilöstö- ja logistiikkakustannuksista. Tämä saattaa johtua organisaatioiden sisäisistä rajoista tai yrityksen laajuisen yhteistyön puuttumisesta. Myös yrityksen rahanjakopolitiikka saattaa ajaa tällaiseen ratkaisuun, koska muutoin ei olisi ratkaisua lainkaan. [Jukic, 2006]



Kuva 13. Itsenäiset paikallisvarastot [Jukic, 2006].

Inmonin metodologian edut ovat kattavuus ja monipuolisuus ja Kimballin metodologian edut ovat nopeus ja yksinkertaisuus. Inmonin lähestymistavassa ER-mallin luominen on ensimmäinen toimenpide. Kimballin lähestymistapa ei vaadi ER-mallia, ainoastaan organisaatiolle tarpeellisten moniulotteisten mallien luomisen. Jos yrityksessä ei tarvita integroitua näkemystä, on Kimballin lähestymistapa nopeampi ja helpompi, mutta jos tarvitaan myös muun tyyppisiä tietorakenteita jatkotarpeita silmälläpitäen, on Inmonin lähestymistapa voimakkaampi. Jukicin [2006] mukaan Inmonin lähestymistapa on perusteltu, koska se mahdollistaa tietovaraston analysoinnin useiden tekijöiden suhteen. Lisäksi, koska ETL-prosessin toteuttaminen vie suurimman osan resursseista, joten mallintamiseen panostaminen on kaukonäköistä. Jukic kumoo myös Kimballin lähestymistapaa kohtaan esitetyn kritiikin yrityksen kokonaisnäkemysten puuttumisesta, perustelemalla, että faktataulut voidaan toteuttaa yrityksenlaajuisesti ja niitä voidaan käyttää samanaikaisesti useassa yrityksen yksikössä. Organisaation onkin valittava se lähestymistapa tai lähestymistapojen yhdistelmä, joka parhaiten sopii heidän organisaatioonsa ja vastaa parhaiten nykyisiin ja tulevaisuuden, ennakoitaviin ja ennakoimattomiin analysointivaatimuksiin. [Jukic, 2006].

3.5. Organisaation tieto ja tietomalli

Tietomallin ja yhteisen käsitteistön merkitystä voi selventää esittämällä ihmisen tapa omaksua tietoa. Andrighetto ja Assogna [2004] pohtivat yhteisen tietämyspohjan merkitystä yhteiselle kommunikoinnille eli käsitteille ja niiden välisille suhteille. Tietämyspohja koostuu heidän mukaan ohjaavasta ontologiasta sekä järjestetystä ja järjestämättömästä tiedosta, jota hallitaan organisaatiossa. Andrighetto ja Assogna korostavat, että tietämys ei koostu pelkästään viesteistä vaan myös niiden vastaanottamisesta ja yhdistämisestä. He toteavat, että henkilö ei voi kopioida jonkun toisen mallia – havaintoa vaan, kun henkilö saa tietoa johonkin asiaan liittyen hänen täytyy muodostaa malli tästä asiasta mielessään. Jokainen meistä muodostaa oman tietämyksen maailmasta ja tämä tieto esitetään sanoina. Sanoihin puolestaan eri ihmiset liittävät erilaisen semantiikan. Koska jokainen meistä ihmisistä on myös yksilö, ei

Andrighetton ja Assognan mukaan voi olla olemassa kahta samanlaista tietämysjoukkoa, kahta samanlaista maailmaa. Heidän mukaan ongelmaksi muodostuu teksteihin ja kuviin liittyvän tietämyksen siirtäminen eteenpäin, jotta tietämys saadaan siirtymään vastaanottajaan mieleen. Kulttuurin ja sosiaalisen kontekstin siirtäminen on haastavaa, joten on kehitettävä metodeja ja tekniikoita kontekstin siirtämiseen ihmiseltä toiselle. Tietovarastointiprojektissa tämä vuorovaikutus on mielestäni erityisen tärkeätä, koska esimerkiksi keskustelukumppanit yleensä edustavat erilaisia toiminta- ja käsitemaailmoja: tietotekniikkaa ja liiketoimintaa.

Vastakkaisen, mutta laajemman näkemyksen yrityksen laajuiseen tietämykseen tuo esille Megill [2005, s. 14] määrittämällä, että informaatio on kadonnut, kun sitä ei voi enää ottaa käyttöön. Esimerkiksi tietoihin ei kyetä liittämään tarvittavaa semantiikkaa. Vaikka Megill kirjoittaakin yrityksen muistista ja sen hallinnasta, tämän voi tarkentaa koskemaan tietovarastointia tiedon säilytyksessä ja yrityksen tietomallia informaation metatiedon säilytyksessä.

Pysyvästi tuhotun tiedon vaikutus on merkittävä yrityksille. Asioita ei välttämättä ole myöskään heitetty pois vaikka ne olisi pitänyt pysyvästi tuhota. Tällöin esimerkiksi dokumenteista voi olla useita versioita ja niistä ei tiedetä mitkä ovat merkittäviä. Informaatio voi myös unohtua, dokumentti voi hukkua massaan tai se voi olla siirretty tuntemattomaan paikkaan ja informaation historia on hävinnyt. Tällaista hiljaisten tietojen ja tapojen katoamista tapahtuu esimerkiksi työntekijöiden jäädessä eläkkeelle. Dokumentoinnin puuttuessa on usein kyse toimintatapojen puutteesta, joskus harvemmin myös salassapitoehdoista. Paras hetki yrityksen tietämyksen kokoamiseen on silloin, kun toiminta on alkamassa ja yrityksen tietämysjärjestelmät ovat osa normaaleja liiketoimintatapoja. Jälkeenpäin dokumentointi on huomattavasti vaikeampaa, koska jälkikäteen asioihin on kyettävä yhdistämään niihin liittyvä semantiikka. Yksi syy tietämyksen katoamiseen voi myös olla tietovälineen katoaminen. Ongelma on kasvanut voimakkaasti elektronisen tietojensäilytyksen yleistyttyä. Ohjelmistoista tehdään keskimäärin uusia versioita kerran vuodessa, ja muutaman vaihtuneen version ja tietojen konvertoinnin jälkeen huomataankin, että joitakin tietoja on kadonnut. Historiallisten tietojen siirtäminen ohjelmien eri versioiden välillä on yksi suurimpia ongelmia tiedonsäilytyksen kannalta. Kadonnut informaatio voidaan palauttaa vain löytämällä se tai luomalla se uudelleen. [Megill, 2005, s. 14-19] Tietovarastolla voikin olla rooli esimerkiksi yrityksen tulevan henkilöstön kouluttamisessa, koska tietovaraston avulla syntyy kuva liiketoiminta-alueeseen kuuluvista keskeisistä tekijöistä.

Informaation säilytysaika on olennainen osa myös tietovarastointia. Megillan [2005, s. 48] mukaan liian tiedon ja liian vähäisen tiedon säilyttäminen ovat yhtä vahingollisia yrityksen tietämyksen säilyttämiselle. On kuitenkin muistettava, että informaation

säilytyspäätös on liiketoimintapäätös, ei yksilön päätös [Megill, 2005, s. 49]. Megillanin [2005, s. 47] mukaan informaation arvon voi tutkia neliosaisella testillä tutkimalla sen

- toiminnalliset tarpeet, mitä informaatiota tarvitaan liiketoimintaa varten
- ylläpitotarpeet, mitä informaatiota tarvitaan ylläpidollisiin tarkoituksiin
- lakivaatimukset, mitä tietoa on säilytettävä lakimääräysten takia
- historialliset tarpeet, mitä tietoa on säilytettävä pysyvästi.

Dokumenttien arvon voi tutkia monin tavoin. Organisaatiolle tärkeiden ihmisten käyttämä informaatio on yleensä myös tärkeää informaatiota. Nämä ihmiset voivat olla millä tahansa organisaation tasolla, mutta kuitenkin yleensä ylimmässä johdossa. Mitä korkeammalla tasolla tieto on, sitä summatumpaa ja tulkitumpaa se on. Organisaation alimman tason tieto on kuitenkin kaiken päätöksenteon perusta. [Megill, 2005, s. 48]

Megill [2005, s. 49-50] toteaa, että informaatio saavuttaa arvon vasta siinä vaiheessa, kun sitä käytetään, toisin sanoen vain silloin, kun se on käyttökelpoista. Yhtenä käyttökelpoisuuden mittarina hän pitää yhteisen ja jaetun informaation käyttöastetta. Analysoimalla informaation käyttäjiä ja käytön laajuutta on mahdollista selvittää informaation käyttötapoja organisaatiossa [Megill, 2005, s. 49-50]. Herrmann [2004] on tutkinut informaation käyttämistä organisaation näkökulmasta ja havainnut, että vain 20 prosentilla tietovarastointihankkeista ei ole ongelmia dokumentoinnin kanssa ja, että se on standardoidusti tuotettu. Lisäksi hän on havainnut, että esimerkiksi yli 80 prosentilla yrityksissä tietovarastoinnissa käytetyt termit ovat vain osittain tai ei lainkaan dokumentoidut. Megill [2005, s. 50-51] toteaa, että yritykselle elintärkeät tiedot ja tallenteet muodostavat kuitenkin yrityksen muistista tärkeän osan. Hän korostaa, että nämä tiedot ovat katastrofeista, kuten tulipaloista selviytymisen edellytyksenä.

Historiallinen informaatio on arvokasta ja sitä tarvitaan yrityksen kehittämiseen, käytäntöjen muuttamiseen, mutta kuitenkin myös yrityksen tapojen säilyttämiseen [Megill, 2005, s. 51]. Megillan mukaan vain ymmärtämällä yrityksen lähtökohdat ja historiallisen taustan on mahdollista tietää, miksi joitakin päätöksiä on matkan varrella tehty. Hän toteaa, että historiallinen informaatio on arvokasta trendien ja esimerkiksi tuotteiden elinkaarien tunnistamisessa. Toisaalta pitää muistaa, että mennyt aika ei määritä välttämättä tulevaisuutta ja kerätty informaatio voi muodostua jopa rasitteeksi.

Informaatio, jonka säilytyksestä määräävät lait ja säädökset, on hyödyllistä lain tai säädöksen ajan, mutta ei pidempään. Useimmat yritykset ajautuvat ongelmiin, kun dokumentit ja informaatio kertautuvat. Hallinnoimattomat tallenteet ja tietojärjestelmät ovat potentiaalisia aikapommeja. Hyvän yrityksen muistin ja informaation hallintaan liittyy muutoksista selviytyminen. Organisaatiot yhdistyvät ja uudelleenorganisoidut. Informaatio ei ole staattista. Toisin sanoen, kun sen käyttö muuttuu, niin myös sen arvo muuttuu. [Megill, 2005, s. 52-53]

Epämuodollinen liiketoimintatietämys on tärkeää yrityksen tehokkuudelle, mutta sitä ei voi järjestelmällisesti käyttää ja se ei ole etu, jonka yritys omistaa. Käsitteelliset kehukset ovat selvästi tarpeellisia yrityksen tietämyksen jäsentämiseen ja hallitsemiseen siten, että se on selkeästi määritelty, hallittu ja se tarjoaa tavan varmistaa, että kehys on käytettävissä ja sitä voidaan tarvittaessa käyttää. Tällaista toimintaa nimitetään yrityksen laajuiseksi mallintamiseksi (engl. enterprise modelling), joka on pelkistetty kuvaus organisaation avaintoiminnoista ja niiden keskinäisistä riippuvuuksista. Yrityksen laajuiseen mallintamiseen liittyvät näkökulmat organisaation sovellusalueista, kuten tavoitteista, rakenteista ja työrooleista, toiminnallisesta alueesta, kuten liiketoimintaprosesseista sekä informatiivisista tiedoista, kuten tietokannan rakenteesta. Avain menestyksellisiin liiketoimintamuutoksiin on yrityksen päätöksentekijöiden tietämys; missä yritys on tällä hetkellä, missä se toivoo olevansa tulevaisuudessa ja suunnitelmat muutoksen aikaansaamiseksi. Yrityksen tietämys on liiketoimintasovelluksen ja näitä tukevan datan yhteenliittymä. [Loucopoulos, 2000]

Loucopoulos [2000] esittelee yrityksen mallintamiseen Delos-nimisen lähestymistavan, joka perustuu kolmeen näkymään. *Organisatorinen näkymä* määrittelee liiketoimintatavoitteet, -toimijat ja säännöt, jotka rajoittavat liiketoimintaa. *Toiminnallinen näkymä* määrittelee, kuinka liiketoiminta toteuttaa tavoitteet päivittäisissä toiminnoissa. *Informatiivinen näkymä* määrittelee tiedon, sen rakenteen ja päivittäistä toimintaa tukevat operaatiot. Jäsennelty liiketoimintatietämys on kehitettävä vaatimusmäärittelytekniikoilla ja käyttämällä sopivia työkaluja ja koko yrityksen kattavaa mallintamista. [Loucopoulos, 2000]

Organisaation tietojen käsittely koostuu ei-teknisestä arkkitehtuurista, vaatimusten määrittelystä ja tietojen analysoinnista. Tietovaraston on heijastettava kaikkia organisaation informaatiotarpeita ja siksi tietovarasto on määriteltävä yleisen, yrityksen informaationäkymän ehdoin, varsinkin mikäli yrityksellä ei ole olemassa yleistä, integroitua näkymää tietoihin, perustelee Calvanese [Calvanese *et al.*, 1999] yrityksen tietomallin tärkeyttä. Myös Moody ja Kortink [2000] toteavat, että yrityksen tietomallin pitää olla keskitetyn tietovaraston perusta.

Yhteinen käsitelmä on täten yrityksen tietomallin peruskivi. Noreen Kendle [2005] kuvaa yrityksen tietomallia integroituna näkymänä koko organisaation tuottamaan ja kuluttamaan tietoon. Tietomallin esittämät tiedot ovat Kendlen mukaan riippumattomia yrityksen järjestelmistä ja se esittää, yhdistää sekä formalisoi yritykselle tärkeät asiat ja niiden säännöt. Jarke [2001] vie organisaation laajuisen mallin vielä pidemmälle ehdottamalla, että kaikki muut mallit määritellään vain näkymiksi suhteessa organisaation laajuiseen malliin. Jarke [2001] perustelee tätä sillä, että tietovarasto on merkittävä sijoitus tiettyyn liiketoimintatarkoitukseen.

Shanksin ja Darken [1999] määrittävät yritystietomallin (engl. corporate data model) seuraavasti:

”organisaation sisäinen, korkean tason malli informaation vaatimuksista”.

Yritystietomalli on abstrakti esitys informaatiovaatimuksista ja itse informaatiosta. Malli esitetään yleensä käyttäen hyväksi käsitteellisen mallintamisen merkintätapoja, kuten ER-mallia [Shanks and Darke, 1999]. Yritystietomallit eivät saa heijastaa tiettyä henkilöä, organisaatorakenteita tai teknologiaa [Shanks and Darke, 1999]. Sen sijaan ne kattavat organisaation toiminnalliset alueet tarjoamalla yhteisen yleisnäkökuvan tietoon. Yritystietomallien käyttämisen oikeutus saadaan esimerkiksi epäjohtomukaisten tietojen poistamisena, tiedon laadun parantumisessa ja tiedon helpommassa integroitavuudessa sekä päällekkäisen tiedon parempana hallintana [Martin, 1982].

Yritystietomallin eräs ongelma on mallin ymmärtämisen vaikeus sekä liiketoiminta-, että tietojärjestelmäihmisille. Yhteisen kommunikoinnin vaikeus johtuu yleisistä käsitteistä, jotka eivät välttämättä liity todelliseen liiketoiminta-alueeseen. Väitösperustelupohjainen suunnitteleminen ja skenaariopohjainen analysointi ovat mekanismeja, joilla sidosryhmien osaaminen saadaan kerättyä. Myös käsitteellisten visualisointi auttaa sidosryhmäläisiä hahmottamaan mallin paremmin. [Shanks and Darke, 1999]

3.5.1. Organisaation ei-tekniinen arkkitehtuuri

Moss ja Atre [2003, s. 65] väittävät, että organisaation ei-tekniinen arkkitehtuuri on kriittinen menestystekijä (engl. critical success factor) rakennettaessa liiketoimintatiedon hallinta- ja päätöksentekoympäristöä. Kriittisillä menestystekijöillä tarkoitetaan sellaisia asioita, joissa organisaation on ehdottomasti onnistuttava. Moss ja Atre [2003, s. 64] ovat esittäneet, että ei-tekniisen infrastruktuurin arvioinnilla pyritään vastaamaan esimerkiksi seuraaviin kysymyksiin:

- Tietojen omistaja, toisin sanoen, kuka vastaa tiedoista ja keiden liiketoimintaihminen on osallistuttava loogisen tietomallin ja liiketoiminnan metatietojen muodostamiseen ja vahvistamiseen?
- Kuka integroi projektin loogisen mallin yrityksen loogiseen malliin?
- Ovatko standardit dokumentoitu ja noudatetaanko niitä?
- Ovatko projektin laatuohjeet riittävän tehokkaita epäpuhtaiden tietojen arviointiin ja puhdistamiseen. Epäpuhtaat tiedot voivat sisältää esimerkiksi arvoja, jotka eivät tietoihin kuuluisi tai tiedoista puuttuvat pakolliset arvot tai tietojen mukana tulee väärää tietoa.

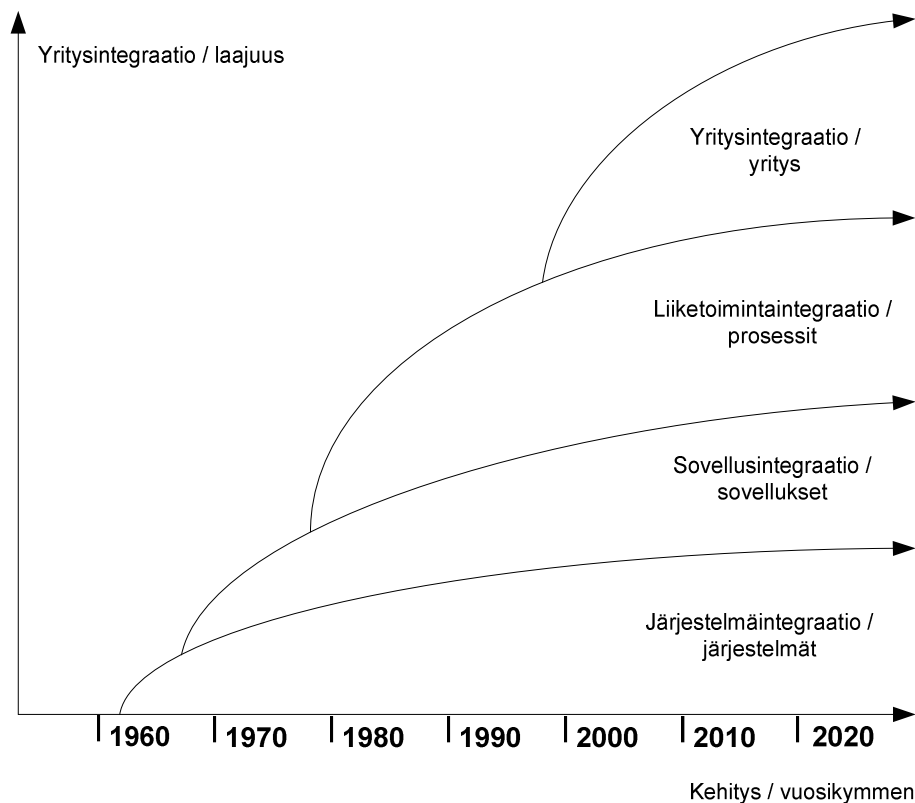
Inmon [2002, s. 64] muistuttaa, että tietovaraston käyttäjäkunta on hyvin heterogeeninen: esimerkiksi suunnittelijoita, kehittäjiä, analyytikkoja. Siksi organisaatiolle on tarpeellista kuvata, mitä tietovarasto pitää sisällään. Tämä Inmonin julkaisuksi nimittämä dokumentaatio esimerkiksi määrittelee tietovaraston, kuvaa sen

lähteet ja käyttötarkoituksen, suhteen lähtöjärjestelmiin ja ohjeet metatietojen käyttämiseen sekä määrittelee vastuut.

3.5.2. Tietointegraatio

Integraation tarkoituksena on tarjota pääsy oikeaan ja tarpeelliseen informaatioon, oikeaan aikaan erilaisissa päätöksentekotilanteissa. Integraation kautta tietojen hallittavuus paranee. Käytettävät resurssit, käytetyt prosessit, organisaation rakenne ja informaatio ovat tärkeimmät rakennuspalikat yrityksen liiketoimintatavoitteiden tietotekniselle tuelle. Kehitys järjestelmäintegraatiosta kohti yritysintegraatiota (engl. *Enterprise Integration*, EI) on esitelty kuvassa 14. Yritysintegraatio on kasvavan kiinnostuksen kohteena yrityksen eri toiminnoissa. Yritysintegraation pääasiallinen tavoite on helpottaa tiedon, hallinnan ja materiaalivirtojen kulkua yli yrityksen sisäisten ja ulkoisten organisaatorajojen eli pyritään yrityksen mutkikkuuden hallintaan. Ajasta ja paikasta riippumaton ongelmienratkaisu ja päätöksenteko johtavat parantuvaan tuottavuuteen ja joustavuuteen yksittäisten yritysten rajoista riippumatta. [Nagarajan *et al.*, 1999]

Järjestelmäintegraatiossa pyritään yhdistämään tietojärjestelmiä suoraan toisiinsa. Sovellusintegraatiossa sovellukset kommunikoivat toistensa kanssa. Liiketoimintaintegraatiossa integraatio keskittyy liiketoimintaprosessien yhdistämiseen.



Kuva 14. Integraation kehittyminen yritysintegraation suuntaan [mukaiillen Nagarajan *et al.*, 1999].

Yrityksen tietomallissa on kyse tietojen ja tietovarastointiprosessien integroinnista ja yhteensovittamisesta. Tietointegraatio tarkoittaa tiedon määritysten ja rakenteiden standardoimista käyttämällä yleistä käsittemallia kaikissa tietolähteissä [Goodhue *et al.*, 1992]. Tietointegraatio varmistaa, että tiedolla on sama merkitys ja käyttötarkoitus eli tulkinta ajasta ja käyttäjästä riippumatta, tekemällä eri järjestelmien ja tietokantojen tietojen esittämisen yhdenmukaiseksi ja loogisesti yhteensopiviksi [Goodhue *et al.*, 1992]. Tietointegraatiota voidaankin pitää organisaation tietomallin perustana. Goodhue ja muut [1992] ovat lisäksi esittäneet yhteisten loogisen tason attribuuttimääritysten ja tietosisällön koodausten käyttämistä organisaation eri osissa.

Goodhue ja muut [1992] ovat tiivistäneet tietointegraation hyödyt parempana kommunikaationa ja koordinoituna, jos organisaation aliyksiköillä on keskinäinen riippuvuussuhde tai niillä on epämääräinen toimintaympäristö. Toisaalta Goodhue ja muut [1992] ovat todenneet että, jos aliyksiköillä on eroavuuksia, on seurauksena yleensä kompromisseja ja enemmän suunnittelukustannuksia sekä byrokratiasta johtuvaa viivettä. He ovat kuitenkin lisänneet, että eroavaisuudet ja liiketoimintaympäristössä tapahtuvat jatkuvat muutokset aiheuttavat vielä enemmän kompromisseja, kuluja ja suunnittelukustannuksia. Nämä haittapuolet on otettava huomioon esimerkiksi arkkitehtuurisilla ratkaisuilla, jotta muuttuviin liiketoimintaperusteisiin voidaan vastata joustavammin.

Kustannusten ja hyödyn käsitteellistäminen tietointegraatiossa on tarpeen, jotta saadaan selville kuinka integrointi vaikuttaa käyttäjien ja toteuttajien mahdollisesti saavutettaviin hyötyihin sekä kulujen syntymiseen. Käsitteellistämisen avulla järjestelmien toteuttamiseen voidaan valita oikeat suunnittelumenetelmät [Goodhue *et al.*, 1992].

Integrointi on inkrementaalinen (asteittain etenevä) prosessi, koska organisaatio ei todennäköisesti pysty sulauttamaan kaikkia tietolähteitä yhdellä kertaa. Toinen inkrementaalista prosessia tukeva peruste on integroinnin jatkuva uusiutuminen uusien tietolähteiden muodossa, jolloin tietosuunnittelijoiden on liitettävä ne järjestelmään. [Trisolini *et al.*, 1999]

Tietointegraatio on keskeinen osa tietovarastointia. Calvinase ja muut [Calvanese *et al.*, 1999] ovat perustelleet tietointegraation keskeistä asemaa tiedonkeruuprosessin tietoa integroivalla luonteella sekä tietovaraston tavoitteella tuottaa yhdenmukainen näkymä (engl. integrated view) tietovaraston tietoon. Operatiivisista järjestelmistä ja sovelluksista kerättyjen tietojen mahdolliset ristiriidat ja toisteisuus ovat ratkaistava, jotta tietovarasto kykenee tarjoamaan integroidun, täsmällisen sekä sovitetun näkymän organisaation tietoon. Toisaalta Calvinase ja muut [Calvanese *et al.*, 1999] ovat myös todenneet tietointegraation ja tietojen sovittamisen tarkoittavan koostettua tietojen keruuprosessia eri lähteistä ja sen saattamista tietovaraston käyttäjien käytettäväksi.

Watson ja Gray [1998, s. 11] ovat todenneet, että tietovaraston informaation on oltava puhdasta, kelvolliseksi osoitettua ja asianmukaisesti summattua. Puhtaalla he tarkoittavat, että samaan tietoon viitataan vain yhdellä tavalla. Tiedot integroidaan, kun ne tuodaan tietovarastoon, jotta tiedoilla ovat samat muodot ja samat mittayksiköt kaikkialla tietovarastossa, vaikka lähteissä muoto ja mittayksiköt voivat vaihdella. Lisäksi Watson ja Gray [1998, s. 11] muistuttavat, että ennen kuin tiedot tallennetaan tietovarastoon on myös varmistettava, että tiedot vastaavat reaali maailmaa ja ne ovat virheettömiä, koska virheellinen tieto säilyy tietovarastossa kauan ja vaikuttaa sitä kautta päätöksentekoon. Toisaalta tapaustutkimusteni perusteella totesin, että tietojen virheellisyys auttaa näyttämään lähtöjärjestelmistä kerättyjen tietojen sisältämät virheet, jolloin ne voidaan tunnistaa ja mahdollisesti korjata operatiivisiin järjestelmiin. Yleensä tietojen korjaaminen kuitenkin tarkoittaisi manuaalista työskentelyä sekä lähtöjärjestelmiin, että tietovarastoon.

Calvanese ja muut [1999] ovat esitelleet erään tavan kuvata käsitteellisesti tietojen integrointia tietovarastoa varten. Käsitteellisen integraation tuloksena syntyvä globaali yritystason näkymä edustaa integroitua näkymää organisaation tietoon silloin, kun integroitua näkymää ei organisaation tietoon ole saatavilla. Oleellista informaation integroinnissa Calvanasen ja muiden [1999] mukaan on organisaatiotason termien käyttäminen lähtöjärjestelmien termien sijaan. He ovat ehdottaneet, että kerättävien tietojen muuntamisprosessi alkaa oleellisia tietoja sisältävien lähteiden tunnistamisella. Seuraavaksi käyttäjien pyynnöt puretaan kyselyiksi, jotta yksittäisistä lähteistä saadaan haluttu tieto esiin. Seuraavaksi tieto tulkitaan ja muunnetaan yleiseen esitysmuotoon. Viimeiseksi toimenpiteeksi he määrittelevät eri lähteiden tietojen yhdistämisen tietovarastoa varten.

3.5.3. Yrityksen tietoarkkitehtuuri

Moss ja Atre [2003, s. 68] sisällyttävät yrityksen tietomallin osaksi yritysarkkitehtuuria. Yritysarkkitehtuurin (engl. enterprise architecture) muodostavat organisaatiota kuvaavat havainnolliset esitykset organisaation todellisista toiminnoista. Näissä arkkitehtuurimalleissa esitetään liiketoiminnan prosessit, prosessien sisältämät toiminnot ja niiden tuottama tieto, joita täydennetään metatiedoilla, kuten standardimääritelmillä, liiketoimintasäännöillä ja menettelytavoilla. [Moss and Atre, 2003, s. 68]

Jokaisella organisaatiolla on yritysarkkitehtuuri, vaikka sitä ei olisikaan dokumentoitu [Moss and Atre, 2003, s. 68]. Jos arkkitehtuuri on dokumentoimatta, organisaation liiketoiminnot ja niiden kohteet eivät todennäköisesti ole yhdenmukaisesti ymmärretty läpi organisaation [Moss and Atre, 2003, s. 68]. Dokumentoinnin päämääränä ovat väärinkäytön ja organisaation laajuuden näkemyksen katoamisen välttäminen [Moss and Atre, 2003, s. 68]. Organisaationlaajuinen näkemys hämärtyy

esimerkiksi luomalla liiketoimintakohteista tarpeettomasti yksilöllisiä ts. järjestelmäkohtaisia määrittelyjä.

Moss ja Atre [2003, 69] jakavat yritysarkkitehtuurin viiteen komponenttiin:

- 1.) liiketoiminnan toimintojen malliin
- 2.) liiketoimintaprosessimalliin
- 3.) liiketoimintatietomalliin
- 4.) sovellusluetteloon
- 5.) metatiedon kuvauskantaan.

Liiketoiminnan toimintamalli (engl. the business function model) kuvaa mitä organisaatio tekee. Malli auttaa järjestämään organisaation hierarkkisen rakenteen liiketoiminta-alueiksi. Vertikaalinen liiketoiminta, kuten esimerkiksi pankin lainanmyöntäminen, vastaa yleensä yhtä suurta mallin liiketoimintaa. [Moss and Atre, 2003, s. 69]

Liiketoimintojen prosessimalli (engl. the business process model) kuvaa, kuinka organisaatio toimii, toisin sanoen, kuinka liiketoimintojen prosessit ovat toteutettu. Liiketoimintaprosessien parantaminen käynnistyy usein liiketoimintatiedonhallintaprojektien ansiosta, ja malli on tällöin oleellinen työkalu. [Moss and Atre, 2003, s. 70]

Liiketoimintatietomallia (engl. the business data model) kutsutaan yleensä yrityksen loogiseksi tietomalliksi (engl. enterprise logical data model) tai yrityksen informaatioarkkitehtuuriksi. Organisaatiossa on vain yksi liiketoiminnan tietomalli ja se kuvaa organisaation varsinaisen liiketoiminnan tiedot — liiketoiminnan eri toimintojen käsitteet ja näiden käsitteiden väliset suhteet. Lisäksi se kuvaa tietoalkiot, jotka näistä käsitteistä talletetaan ja liiketoimintasäännöt, jotka ohjaavat käsitteiden toimintaa. Tieto-oliot ja tietoalkiot ovat tietomallissa ainutlaatuisia, vaikka niitä säilöttäisiin useassa tietokannassa, koska ne esiintyvät todellisuudessa vain kerran. Liiketoimintatiedonhallinnan ja päätöskijärjestelmän kehityksessä tärkeimmät eitekniikan arkkitehtuurin komponentit ovat liiketoimintatietomalli ja metatiedon kuvauskanta. [Moss and Atre, 2003, s. 70]

Sovellusluettelo pitää kirjaa liiketoimintojen, liiketoimintaprosessien ja liiketoimintatietojen fyysisistä toteutuksista. Se kertoo arkkitehtuurin osien (kuten tietokantojen) sijainnit ja suhteet teknisessä arkkitehtuurissa. [Moss and Atre, 2003, s. 70]

Metatiedon kuvauskantaan kerätään liiketoiminnan analysoinnin ja suunnittelun aikana täydentävää sekä kuvailevaa informaatiota edellä kuvatuista malleista. Näitä tietoja ovat esimerkiksi tietojen semantiikan määrittäminen sekä omistajat, sarakkeiden sallitut arvot, taulujen ja raporttien nimet sekä kuvaukset. [Moss and Atre, 2003, s. 71]

3.5.4. Zachman Framework

Organisaation tietomalli on osa yrityksen laajuista kehystä, jonka avulla järjestelmät voivat toimia yhteistyössä tehokkaammin [Silverston, 2001, s. 3-4]. Tietovarastoinnin voi sijoittaa tietojärjestelmien arkkitehtuurikuvauksissa monelle käsitteelliselle tasolle. Zachman Framework on malli tietojärjestelmien arkkitehtuurin laadintaan [Zachman, 1987]. Zachmanin malli tarjoaa kokonaisvaltaisen perspektiivin tietojärjestelmien arkkitehtuuriin, suunnitteluun, rakentamiseen, käyttämiseen ja ylläpitämiseen. Se kuvaa järjestelmän kehityksen elinkaaren eri roolien ja näkökulman mukaan [Sowa and Zachman, 1992]. Näkökulma on kokoelma erilaisia strategioita järjestelmän arkkitehtuurista, joihin voidaan vastata tarkoitusta selventävien kysymysten avulla. Kysymys ”mitä?” kuvaa mallinnettavaa tietoa, kysymys ”kuinka?” kuvaa organisaation toimintoja ja prosesseja, kysymys ”missä?” kuvaa sijaintia ja verkkoa, kysymys ”kuka?” kuvaa organisaation ihmisiä, kysymys ”milloin?” kuvaa tapahtumien ajanhetkeä ja aikataulutusta, kysymys ”miksi?” kuvaa motivaatiota, tavoitteita sekä sääntöjä [Zachman, 1987; Sowa and Zachman, 1992]. Roolien avulla arkkitehtuurin eri tasoja voidaan tarkentaa. Esimerkiksi suunnittelijan roolin avulla syntyy järjestelmämalli, jolla vastataan esimerkiksi ”miksi”-kysymykseen liiketoimintasääntömallin avulla. Zachmanin malli voidaan kuvata taulukkona, jossa vaakasuorat tasot kuvaavat rooleja sekä toimijoita ja pystysuorat akselit kuvaavat näkökulmaa.

Vaikka Zachmanin mallia voidaankin soveltaa koko tietovaraston suunnitteluun, esittelen vain pienen osan mallin soveltuvuudesta tietovarastointihankkeen järjestelmäkehitykseen. Tässä tutkimuksessa on kyse omistajien ja suunnittelijoiden tasoista sekä tietojen ja toimintojen akseleista. Zachmanin mallin mukaan omistajan rooli kuvaa liiketoiminnan mallia ja määrittelee liiketoimintatermein käyttö- ja menettelytapoja, joiden lopputuloksena on käsitteellinen organisaation laajuinen liiketoimintamalli [Sowa and Zachman, 1992]. Esimerkiksi omistajan tasolla tietoa voi kuvata ER-mallin avulla ja toimintoja liiketoimintaprosessimallina [Zachman, 1987]. Suunnittelijan työ näkyy loogisena järjestelmämallina, joka mallintaa tietoa loogisten tietomallien ja vaatimusten avulla ja toimintoja tietovirtakaavioiden avulla [Sowa and Zachman, 1992]. Samaan tapaan arkkitehtuurin kuvausta voi soveltaa eri roolien mukaan tietovarastoinnin näkökulmasta. Esimerkiksi omistajan perspektiivillä organisaation keskeiset liiketoimintakäsitteet kuvataan käsitekaaviona, käsitteisiin liittyvät liiketoimintaprosessit kuvataan prosessimalleina ETL-prosessointia varten. Muina dimensioina ovat esimerkiksi projekti- ja asiakasorganisaation roolien ja vastuuden määrittelyt kuhunkin prosessiin, liiketoimintayksiköiden sijoittuminen kuhunkin prosessiin, prosessien keskeisten riippuvuuksien määrittäminen ajallisesti ja toimintatavoista ja standardeista sopiminen vaatimustenhallintaa silmällä pitäen. Lisäksi esimerkiksi Jones ja Song [2005] ovat kehittäneet dimensionaaliseen mallintamiseen malleja ja toimintatapoja, jotka hyödyntävät Zachmanin mallia.

3.5.5. Tietomallin hyödyt

Yrityksen tietomalli auttaa Longin [1998] mukaan kohtaamaan nykyiset ja tulevat haasteet tietoarkkitehtuurin luonnissa. Hän esittelee hyötyinä:

- yritystason perspektiivin
- organisaation strategisen informaatiotarpeen määrittelemisen
- aikasidonnaisen tietovaraston perustiedon muodostamisen
- tietovaraston julkaisustrategian tukemisen
- liiketoimintatiedon johtamisen käynnistymisen.

Yritystason perspektiivin avulla johto voi Longin [1998] mukaan muodostaa kokonaisnäkemyksen yrityksen tietoteknisestä tasosta ja sen muodostamasta liiketoiminnan tuesta. Samoin tietojärjestelmien kanssa toimivat työntekijät hahmottavat paremmin kokonaisuuden ja voivat tarjota ratkaisuja paremman informaation saavuttamiseen [Long, 1998; Silverston, 2001, s. 408-410]. Long [1998] jatkaa, että määrittelemällä strategisen informaatiotarpeen yrityksen liiketoiminnasta poimitaan kriittiset informaatio-osat, joiden avulla dataa voidaan jatkojalostaa tiedoksi. Kun tietovaraston perusta on muodostettu atomisesti, eli perustiedosta, tietovarasto on toiminnaltaan muista järjestelmistä riippumaton järjestelmä. Se on silloin itsenäinen myös sisältämänsä datan suhteen. Tietomalli tukee tietovaraston julkaisustrategiaa auttamalla valitsemaan tärkeimmät liiketoiminnan osa-alueet toteutettaviksi [Long, 1998]. Lisäksi liiketoimintatiedon johtaminen käynnistyy, kun mallin avulla tietoalueet ja näihin liittyvät liiketoiminta-alueet voidaan jaotella ja niille voidaan osoittaa vastuuhenkilöt sekä avainhenkilöt, jotka ymmärtävät järjestelmän merkityksen liiketoiminnalle [Long, 1998].

Silverston [2001, s. 418-419] on todennut yrityksen tietomallin merkittävimpanä hyötynä liiketoimintaongelmien ratkaisukyvyin. Tällöin liiketoimintaihmisillä on käytössään graafisesti kuvattu selkeä näkymä kaikkeen informaatioon, jolloin informaatiota voidaan esimerkiksi organisoida uudelleen. Lisähyötynä Silverston on maininnut mallin helpottavan toisteisten tietojen sekä integroimista vaativien järjestelmien tunnistamista. Mallin avulla voidaan esimerkiksi havaita, että kaksi tai useampi järjestelmä käyttää ja ylläpitää samaa informaatiota. Kendle [2005] lisää, että tiedon laatu voi parantua, koska tietomalli paljastaa tiedon eroavaisuudet toistuvasta datasta. Kendlen mukaan mallia käytetään myös selvittämään jaetun tiedon omistajuus yli organisaatio- ja liiketoimintarajojen. Hän jatkaa, että järjestelmä on laajennettava, koska yrityksen tietomalli on riippumaton teknologiasta ja malli perustuu strategiseen liiketoimintanäkymään. Tietomalli sisältää ulkoisen näkymän, joka lisää yrityksen kykyä teollisuudenalan sisäiseen tietointegraatioon jakamalla yhteisiä käsitteitä.

Tietomallin avulla voidaan tukea paketoitujen ohjelmien integrointia ja linkaarta: ostoja, suunnittelua ja toteutusta, koska tietomalli näyttää yrityksen tietovirrat ja

integrintikohdat [Kendle, 2005; Silverston 2001, s. 408-410]. Riippumaton tietomalli mahdollistaa lisäksi strategisen järjestelmäsuunnittelun, koska kaikki järjestelmät ovat kytkettyinä yrityksen tietomalliin [Kendle, 2005].

Jarke ja hänen tutkijakollegansa [1998] muistuttavat, että laatuorientoitunut metatietojen hallinta vaatii koko tietoketjun läpikäymistä: yrityksen operatiivisten yksiköiden toiminnoista tietovaraston tietojen analysointiin. Tämä vuorostaan merkitsee heidän mukaansa, että laatuorientoituneen arkkitehtuurin keskeinen ominaisuus on yksikäsitteinen yrityksen laajuinen tietomalli. He muistuttavat, että muodostamalla mallin vaiheittaisesti, sen käyttö ei ole riskialtista, eikä liian kallis ponnistus. Yrityksen tietomalli voi syntyä esimerkiksi vaiheittaisena toteutuksena lähteiden integroinnin tai liiketoimintaprosessien kartoituksen yhteydessä.

Organisaation laajuisen mallin ja helposti ymmärrettävän käsitteellisen näkökulman hyötyinä Jarke [2001] näkee:

- järjestelmäriippumattoman lähestymistavan informaation integrointiin
- käyttäjärajapinnan korkean käsitteellisyystason
- inkrementaalisen lähestymistavan
- dokumentaation
- ylläpidon

Arkkitehtuurin korkea käsitteellinen taso mahdollistaa järjestelmäriippumattoman määrittelyn lähteiden sekä lähteiden ja organisaation laajuisen mallin välillä. Kun käsitteelliset mallit esitellään graafisesti, käsitteellinen mallintaminen tarjoaa korkeamman tason suunnitteluun ja lähentää käyttäjää sekä järjestelmää. Käsitteellinen malli voidaan luoda myös asteittain, inkrementaalisesti. Tällöin uusia lähteitä tai elementtejä voidaan lisätä myöhemmin. Erityisesti suurten organisaatioiden tieto on hajallaan dokumentoituna ja niihin saattaa olla vaikea päästä käsiksi. Lisäksi dokumentointi ei välttämättä noudata yleisiä standardeja. Käsitteellinen mallintaminen tarjoaakin tietojen integrointiin pohjan yritykseen varastoituun tietoon ja se voidaan nähdä muodollisena ja täsmällisenä määrittelynä tietolähteiden suhteiden välittäjänä. Suunnitteluvaiheen hyödyt toteutuvat myös ylläpitovaiheessa ts. lähteet vaihtuvat, joten suunnittelu ei lopu koskaan. [Jarke, 2001]

Organisaation tietomalli tarjoaa myös organisaation sovelluskehitykseen uusia näkökulmia. Yhtenä käyttötarkoituksena Silverston [2001, s. 408-410] on esitellyt nykyisten loogisten tietomallien validoinnin ja jatkokehittämisen organisaation tietomallin avulla. Hän on lisäksi ehdottanut, että organisaation tietomalli toimii luontevana lähtökohtana loogisten tietomallien kehittämiselle.

Integroitu yrityksen käsitelmä parantaa tietovarastointisovelluksen kehittämistä. Käsitelmän avulla sovelluskehittäjät ja liiketoimintaihmiset jakavat yhteisen ja merkityksen antavan tulkinnan yrityksessä olevaan tietoon. Yrityksen käsitelmällä on

merkittävä rooli tietämyksenhallintaprosessissa, koska se mahdollistaa yhtenäisen yrityksen laajuisen tietämyskannan rakentamisen edistämällä tietämyksen löytämistä, esittämistä ja levittämistä kaikkialla yrityksessä. [Calvanese *et al.*, 2006]

4. Tietovaraston kehittämisprosessi

Inmon *et al.* [1997, s. 315] kuvaavat tietovaraston tilaajien visiota prosessin lopputuloksesta täydellisesti integroituna yrityksen yhteisenä tietovarastona, jonka data on puhdasta, suhteessa jokaiseen toiseen elementtiin, ajan tasalla ja vastaa sekunneissa kaikkiin kyselyihin. Edellä kuvattuun lopputulokseen voidaan heidän mukaansa päästä, mikäli aloitetaan pienestä kokonaisuudesta ja edetään inkrementaalisiin askeleihin paikallisvarasto kerrallaan.

4.1. Menestystekijät tietovaraston luomiseen

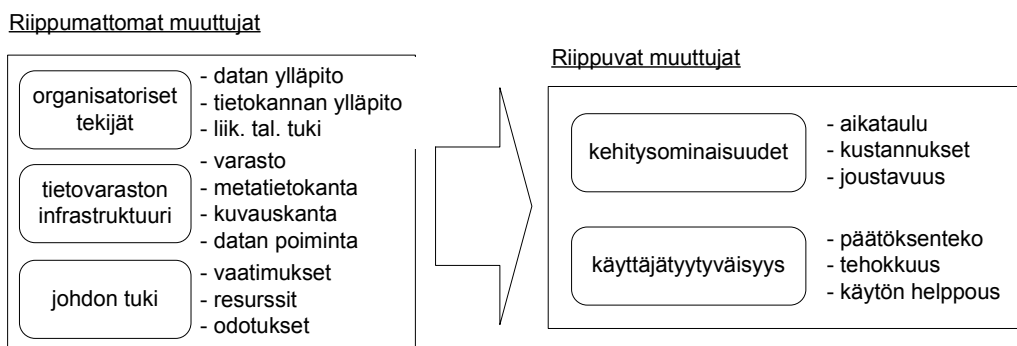
Tietovarastoteutuksen onnistuminen on monesta seikasta riippuvainen. Lisäksi on arvioitava itse tietovaraston onnistumista. Auttoiko se ratkaisemaan liiketoiminnan haasteita?

Tietovaraston luominen on koko organisaation asia. Erityisesti ylimmän johdon tuki on tärkeä, jotta tietovarastoa voidaan hyödyntää tehokkaasti [Kimball and Ross, 2002, s. 368].

Tietovaraston rakentamisessa on tärkeää, että kehittäjät ovat läheisessä yhteistyössä liiketoimintayksiköiden kanssa [Kimball and Ross, 2002, s. 367]. Kimball ja Ross toteavat liiketoimintaihmissen olevan aina oikeassa. Liiketoimintaihmiset, toisin sanoen tiedon loppukäyttäjät, määrittelevätkin tiedon käytettävyyden: vastaako se tiedon tarpeeseen ja soveltuuko se ratkaisemaan liiketoimintaongelman.

Kimball ja Ross [2002, s. 367] toteavat operatiivisten järjestelmien olevan merkittävässä roolissa tietovaraston tiedon keruussa ja puhdistamisessa. He korostavat tätä erityisesti organisaation näkökulmasta, koska tietovaraston keskeinen rooli on ymmärrettävä.

McFadden [1996] on kartoittanut joitakin tietovaraston kehittämiseen liittyviä tekijöitä ja niiden vaikutusta (kuva 15). Näiden tekijöiden avulla hän kuvaa käyttäjätyytyväisyyden ja tietovaraston kehittämisen piirteiden riippuvuutta tietovaraston infrastruktuurista, johdon tuesta sekä organisaation vaikutuksesta. McFaddenin käyttäjätyytyväisyyden mittareina ovat käytön laajuus päätöksenteossa, tehokkuus vasteajan perusteella, datan ja tiedon laatu sekä käytön helppous. Toinen komponentti mallissa on kehityspiirteet, joita ovat aikataulussa pysyminen, kustannukset sekä järjestelmän mukautuvuus uusiin vaatimuksiin.



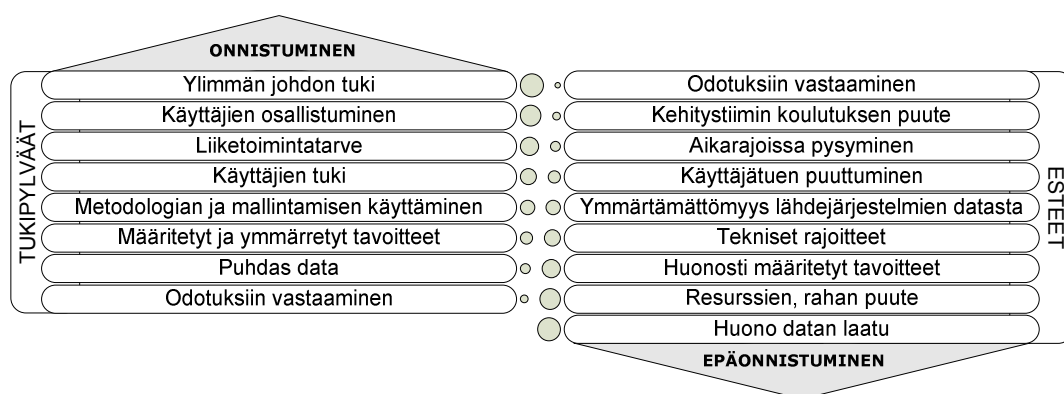
Kuva 15. McFaddenin [1996] kartoittamia tietovaraston menestystekijöitä.

Hyvän tietovaraston ominaisuuksiin kuuluu myös hyvä suunnittelu. Inmon ja muut [1997, s. 324] tunnistavat suunnittelun tärkeimmäksi tavoitteeksi tukea loppukäyttäjän pääsyä tietovarastoon ja sen tietoihin. Lopputulosta voi arvioida esimerkiksi seuraavien Inmonin ja muiden [1997, s. 325] esittämien tunnistettavien piirteiden avulla, joissa näkyy käyttäjän uusien asioiden omaksumistapa. Tietovaraston rakenteen on oltava loppukäyttäjälle helposti tulkittava ja selkeä. Lisäksi käyttäjät tunnistavat liiketoimintansa tuottamat tiedot ja pystyvät käyttämään tietovaraston tietoja ilman merkittävää ulkopuolista apua, jolloin tietovaraston käytön pitää olla helppoa. Seuraavaksi käyttäjille näkyy kyselyiden suoritus aika, joten vasteajan on oltava riittävän nopea. Onnistuneen lopputuloksen tunnistaa käyttäjien halusta laajentaa tietovarastoa.

Menestystekijöitä voidaan tarkastella myös toisesta, päinvastaisesta näkökulmasta, eli mitkä tekijät voivat aiheuttaa tietovarastoprojektin epäonnistumisen. Epäonnistumisen syyt voidaan luokitella esimerkiksi suunnitteluvirheisiin, teknisiin tekijöihin, menetelmätapojen ongelmiin ja sosiaalisteknisiin tekijöihin [Vassiliadis, 2000]. Tekninen virhe on esimerkiksi väärin komponenttien valinta ja organisaation sopimusrajojen rikkoutuminen voidaan kategorioida sosiaalistekniseksi ongelmaksi [Vassiliadis, 2000]. Tämän tutkimuksen kannalta kiinnostavimmat ja oleellisimmat alueet ovat kuitenkin suunnittelu- ja menetelmätapatekijöihin liittyvät ongelmakategoriat. Suunnitteluongelmien kenttään sijoittuva metatiedon hallinnan puuttuminen johtuu suurimmaksi osaksi siitä, että ei ole olemassa standardia tai laajasti hyväksyttyä metatiedonhallintateknologiaa ja tietovaraston suunnittelumetodologioita. Toimittajakohtaisia, eli tietyn yrityksen erityistarpeet huomioivia ratkaisuja löytyy yrityksiltä, mutta avoimia käytännön suunnittelumenetelmiä on Vassiliadis [2002] tutkimuksen mukaan tarjolla vähän tai ne keskittyvät tekniseen suunnitteluun ja integrointiin. Perinteisten tietojärjestelmien ongelmien lisäksi menetelmätavoista löytyvät tietovarastoprojektin kannalta ongelmallisimmat alueet: käyttäjyhteisön puuttuminen suunnittelusta ja sidosryhmien perehdytyksen puute [Vassiliadis, 2000]. Menetelmätapojen ja sosiaalisteknisten kategorioiden ongelmat ovat Vassiliadis

[2000] havaintojen mukaan tutkimusyhteisössä vähiten tutkituimpia alueita tietovarastoinnin ongelmakentässä.

Watson ja Gray [1998, s. 195] esittelevät Watsonin ja Haleyn [1997] tutkimuksen tuloksia tietovarastoinnin menestystekijöistä ja epäonnistumisiin johtaneista syistä. Watson ja Haley selvittivät mm. tietojen laadun, resurssien, käyttäjien osallistumisen ja teknologian osaa tietovarastointitoteutusten onnistumisiin sekä epäonnistumisiin [Wixom and Watson, 2001]. Tutkimuksen lopputulokset on esitetty kuvassa 16, jossa eri tekijät on järjestetty painoarvon mukaan nousevaan (onnistumiset) ja laskevaan (epäonnistumiset) järjestykseen [Watson and Gray, 1998, s. 195].

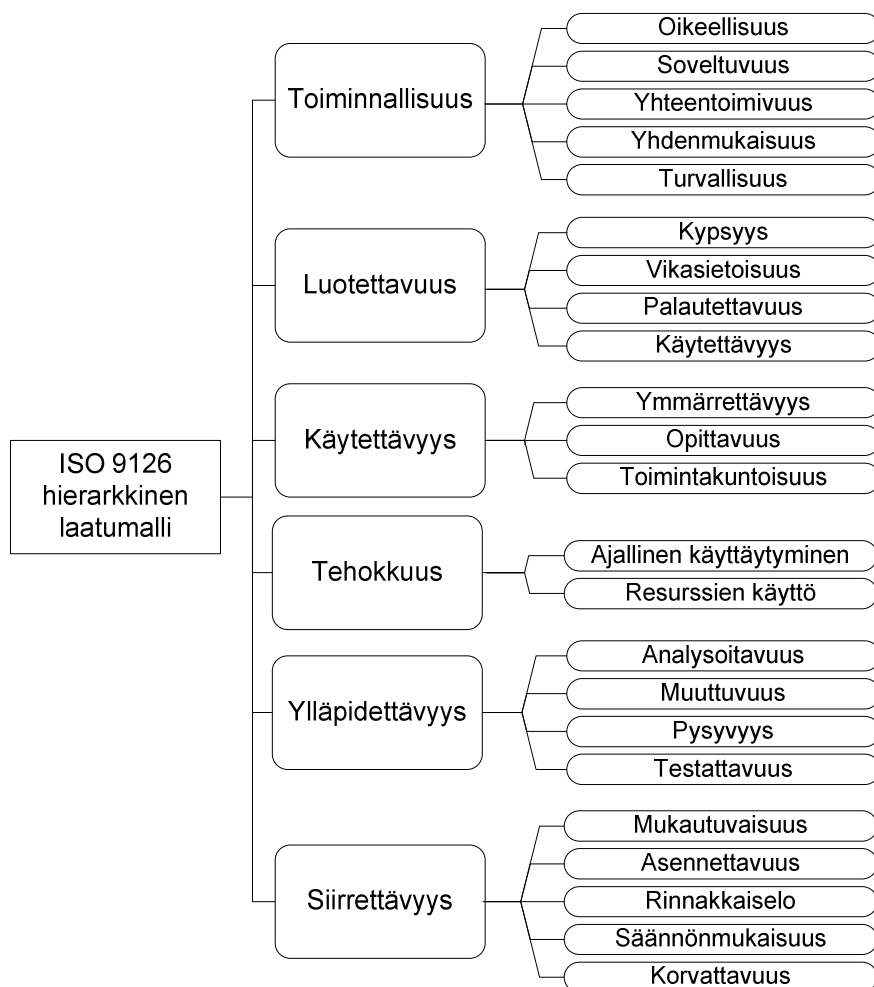


Kuva 16. Tietovarastoinnin kriittiset menestystekijät ja esteet vaikutusjärjestyksessä. Nuolen suunta ja pallojen koko osoittavat vaikutuksien suuruudet.

Ylimmän johdon tuki on merkittävin yksittäinen tekijä projektin onnistumisen kannalta. Suurimmat esteet projektien onnistumiselle olivat tietojen laatuun liittyvät seikat. [Watson and Gray, 1998, s. 195-196]

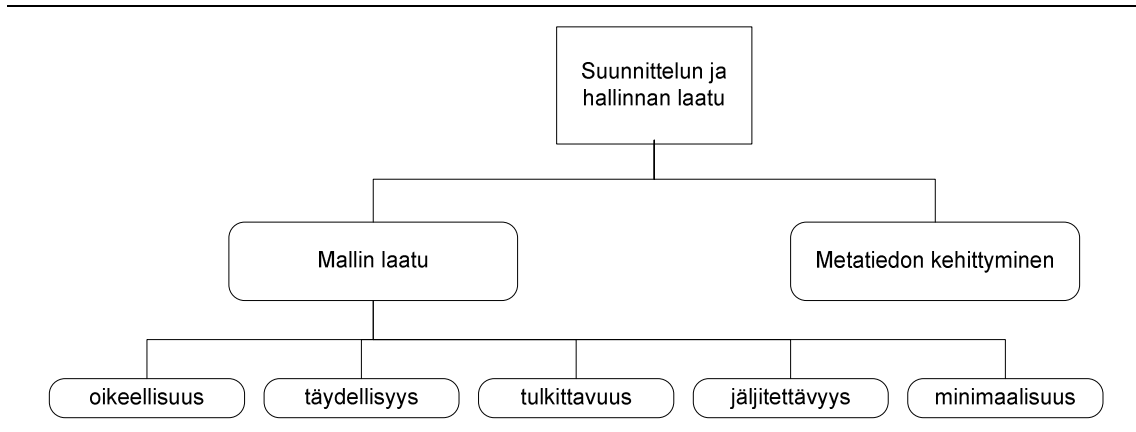
4.1.1. Tietovarastoprosessin laadulliset tavoitteet

Tietovarastoprosessin laadullisia tavoitteita voidaan verrata ohjelmistosuunnittelussa käytettyihin laatumalleihin. Yksi tällainen malli on International Organization for Standardization-järjestön ISO 9126-standardi [Jarke *et al.*, 1998]. Sen laatuun vaikuttavat ominaisuudet on esitelty kuvassa 17.



Kuva 17. ISO 9126-standardin laatutekijät esitettynä lohkokaaavana.

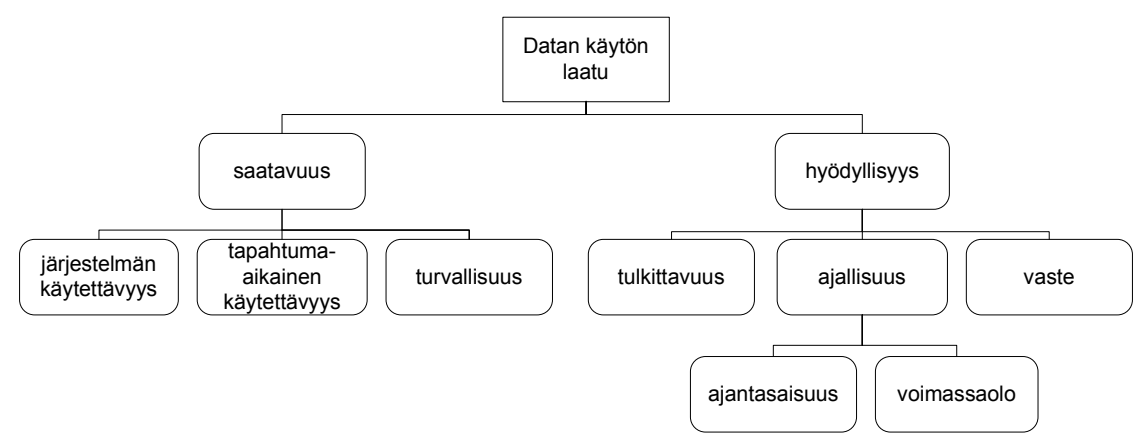
Tietovarasto ympäristön erilaisten sidosryhmien roolien perusteella voidaan tutkia, mitkä edellä esitellyn ISO-standardin laatutekijät kiinnostavat luvussa 2 esiteltyjä rooleja. Päätöksentekijä on yleensä kiinnostunut tallennettujen tietojen laadusta, ajallisuudesta ja raportointityökalun helppokäyttöisyydestä. Tietovaraston ylläpitäjä tarvitsee työnsä tueksi virheraportointia, pääsyn metatietoon ja osaamista tietojen aikasidonnaisuudesta esimerkiksi tietojen muutosten seuraamiseksi. Tietovaraston suunnittelijan on mitattava tietovaraston mallin ja metatiedon laatua. Tietovaraston ohjelmoijat puolestaan hyödyntävät työnsä tukena ohjelmistototeutuksen standardeja. Metatietojen raportointi auttaa heitä välttämään toteutustyössään määrittelykokonaisuuteen liittyviä virheitä. Tämän analyysin perusteella Jarke *et al.* toteavat, että erilaiset roolit määrittävät erilaiset laatutarpeet. Tietovaraston suunnittelija ja ylläpitäjä on kiinnostunut kuvassa 18 kuvatuista laatuattribuuteista. Tietovaraston kehittäjiä ja ohjelmoijia koskevat ohjelmistokehityksen laatuvaatimukset ja näihin Jarke ja muut soveltavat edellä kuvattua ISO-standardia. Tämän tutkimuksen kannalta keskeisimmät laatu kriteerit eli päätöksentekijän laatuattribuutit kuvaan tarkemmin kappaleessa 5.5. [Jarke *et al.*, 1998]



Kuva 18. Tietovaraston suunnittelun ja hallinnan laadun dimensiot esitettynä hierarkkisenä kaaviona.

4.1.2. Tietovaraston laadun mittarit

Tietovarastointiprosessin tärkein osa ovat tietojen käyttömahdollisuudet, toteaa Jarke ja muut [1998]. Jarke ja muut jakavat tietovaraston laatuksiterit loppukäyttäjän näkökulmasta kuvassa 19 kahteen osaan: saatavuuteen (engl. accessibility) ja hyödyllisyyteen. Saatavuudella he tarkoittavat pääsyä järjestelmän sisältämiin tietoihin. Saatavuuteen vaikuttavat käytön valtuuskriteerit, järjestelmän päälläoloaika ja tietojen lukumahdollisuutta, esimerkiksi päivitysvaiheessa. Hyödyllisyys puolestaan tarkoittaa tietojen hyödynnettävyyttä sille tarkoitetussa tehtävässä eli ajantasaisista ja ajallisesti voimassaolevia tietoja, vasteajaltaan riittävän nopeata sekä tietovaraston mallin ymmärrettävyys ja helppokäyttöisyys.



Kuva 19. Tietojen käytön laadun ulottuvuuksia esitettynä hierarkkisenä kaaviona.

4.2. Tietovaraston kehitysprosessi

Tietovarastoinnin tehtävät koostuvat suunnittelu-, rakentamis-, käyttö- ja ylläpitoprosesseista [Staudt *et al.*, 1999b]. Tietovarastointiratkaisu toteutetaan loppujen

lopuksi teknisinä toimenpiteinä. Tietovaraston toteutukseen tarvitaan prosessi, jonka lopputuloksena ovat alkuperäiset järjestelmälle asetetut tavoitteet ja hyvän järjestelmän vaatimukset täyttävä tietojärjestelmä. Tietovarastointiratkaisun toteutusprosessia voi verrata ohjelmistoratkaisun prosessiin ja havaita, että molempien prosessien laadulliset tavoitteet ovat hyvin lähellä toisiaan.

Böhnlein ja Ende [2000] toteavat, että monet tietovarastointiprojektit epäonnistuvat, koska käytössä ei ole suunnittelustrategiaa tai mallintamisstandardeja. He muistuttavat, että tietovarastointiratkaisun kehittäminen eroaa perinteisen järjestelmän kehittämisestä ja tällöin lähtöjärjestelmien tietorakenteiden tarkastelun lisäksi on huomioitava myös yhtiön tavoitteet ja strategia.

Yksinkertaisimmillaan tietovaraston suunnitteluprosessi koostuu neljästä vaiheesta, joista jokaisen vaiheen tuotokset toimivat seuraavan vaiheen lähtökohtina [Rizzi *et al.*, 2006].

- 1.) Liiketoimintavaatimusten analysoimisesta, jonka lopputuloksena ovat tietovaraston vaatimukset.
- 2.) Käsitteellisestä suunnittelusta, jonka lopputuloksena on käsitteellinen malli.
- 3.) Loogisesta, esityksellisestä suunnittelusta, jonka lopputuloksena on looginen tietomalli.
- 4.) Fyysisestä suunnittelusta, jonka lopputuloksena on fyysinen malli.

Rizzin *et al.* [2006] esittelemät suunnitteluvaiheet vastaavat Navathen [1992] esittelemää tietokannan suunnitteluprosessia. Mielestäni tietovaraston suunnitteluprosessi on kuitenkin paljon laajempi ja monimuotoisempi kokonaisuus kuin Rizzi *et al.* esittelevät.

Eritasoisia ja -tyyppisiä prosessimalleja tietovaraston kehitysprosessiksi on useita, mutta vakiintunutta tapaa siihen ei kirjallisuudessa ole esitetty. Moss ja Atre [2003] ovat esittäneet kehitysprosessiksi erään kokonaisvaltaisen ja mielestäni tietovaraston eri näkökohdat hyvin huomioivan lähestymistavan. Se koostuu kuudesta keskeisestä vaiheesta ja on luonteeltaan iteratiivinen, ketterä ja mahdollistaa kehityksen hallinnan usean aliprojektin avulla [Moss and Atre, 2003, s. 5-6, 9]. Vaiheet suoritetaan perätysten, mutta jokaisen vaiheen sisällä suoritettavat toimenpiteet voidaan suorittaa samanaikaisesti [Moss and Atre, 2003, s. 14]. Prosessin päävaiheet ovat (kuva 20)

- 1.) liiketoiminnallisen oikeutuksen määrittäminen
- 2.) strategisen ja taktisten suunnitelmien laatiminen projektin läpiviemiseksi
- 3.) liiketoiminta-analysoinnin toteuttaminen keräämällä liiketoimintavaatimukset liiketoimintaongelman ratkaisemiseksi
- 4.) suunnittelu liiketoimintaongelmalähtöisesti
- 5.) rakennusvaihe, joka takaa investoinnin kannattavuuden määritetyssä ajassa

6.) käyttöönotto toteuttamalla tai hankkimalla ratkaisu, josta arvioidaan saavutettiin asetetut investointitavoitteet.

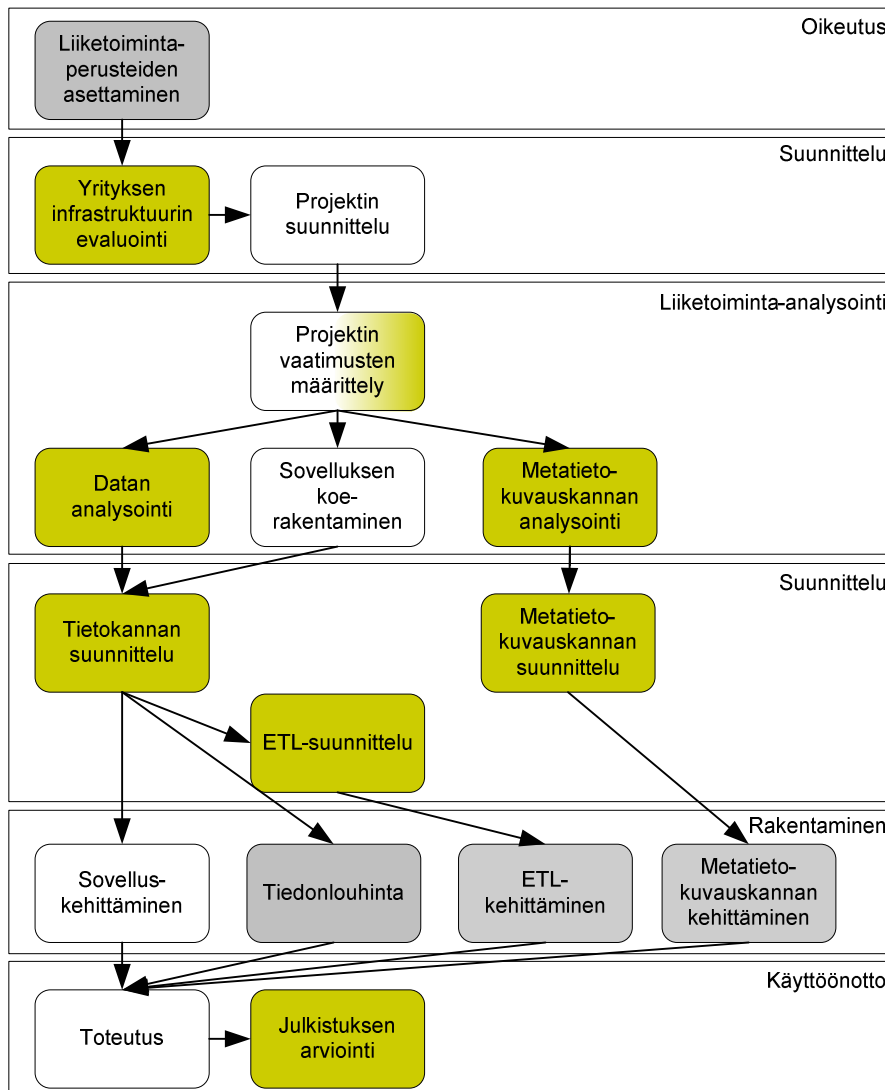
Tämä kehitysprosessi eroaa monista kilpailijoistaan painottamalla organisaation rajat ylittävää (engl. cross-organizational) lähestymistapaa. Atre ja Moss [2003, s. 9] muistuttavat, että organisaation rajat ylittävä kehittäminen vaatii kulttuurimuutoksen organisaation toimintaan. Kehitysprosessin tavoitteena ovat organisaation informaation keskittäminen, integrointi ja eheyttäminen sekä liiketoimintaprosessien virtaviivaistaminen ja päätöksentekojärjestelmien kiinteä kytkeminen elimelliseksi osaksi liiketoimintaa [Moss and Atre, 2003, s. 8-9].

Tämän tutkimuksen kiinnostusalueena ovat kehitysprosessin organisaation rajat ylittävät osat, jotka on väritetty kuvassa 20 joko keltaisella tai harmaalla [Moss and Atre, 2003, s. 16]. Jaottelin kehitysprosessin väritetyt vaiheet vielä tavoitteiden asetteluun, suunnitteluun ja kehittämiseen. Näistä vain suunnittelu kuuluu tämän tutkimuksen kohdealueeseen ja nämä vaiheet ovat väritetty kuvassa 20 keltaisella. Lisäksi mielestäni projektin vaatimusten määrittely kuuluu organisaation rajat ylittäviin vaiheisiin. Näin ollen jäljelle jäävät Mossin ja Atren [2003, s. 11-16] kehitysprosessin seuraavat vaiheet:

- Yrityksen infrastruktuurin evaluointi, jonka tavoitteena on määrittää tekninen ja ei-tekninen kehitysympäristö, kuten tietokannan valitseminen ja käytännöistä sopiminen.
- Projektin vaatimusten määrittäminen on projektin vaativin vaihe, jonka tavoitteena on määrittää esimerkiksi liiketoimintavaatimukset sekä rajoittaa halutut vaatimukset hallittaviksi ja toimituskelpoisiksi osiksi.
- Tietojen analysointi-vaihe vie projektin aikataulusta suuren osan. Lähtöjärjestelmien tietojen laatu muodostaa merkittävän uhkan ja organisaation työntekijöiden vanhojen tapojen muuttaminen on vaikeaa.
- Metatietokuvauskannan analysointi on vaihe, jossa tekniset metatiedot kytketään liiketoiminnan metatietoihin. Yleensä käyttöön otetaan jokin mallinnustyökalu.
- Tietokannan suunnitteluvaiheen tavoitteena on saada tietokantakuvaukset vastaamaan liiketoiminnan käyttövaatimuksia.
- ETL-suunnittelun avulla hallitaan projektin monimutkaisimman prosessin, ETL-prosessin suunnittelua, huomioimalla huono lähtötietojen laatu sekä prosessin ajamiseen käytettävissä oleva aikaikkuna.
- Metatietokuvauskannan suunnittelussa on otettava huomioon asiat, joita käytettävissä olevat työkalut eivät tue. Esimerkiksi kaikki loogiseen tietomalliin kuvatut piirteet on saatava talletettua kuvauskantaan.

- Julkistuksen arviointi on tärkeä vaihe, koska sen avulla voidaan oppia aikaisemmasta toiminnasta ja vaiheen tavoitteena onkin työkalujen, tekniikoiden, ohjeiden sekä prosessin arviointi, sopeuttaminen ja mahdollisesti jopa hylkääminen.

Näissä kaikissa edellä mainituissa kehitysprosessin vaiheissa pitäisi olla mukana organisaation liiketoiminnan osaajia koko organisaatiosta sekä projektin toteuttavasta projektiryhmittymästä. [Moss and Atre, 2003]



Kuva 20. Moss:n ja Atren esittelemä tietovaraston kehitysprosessi (mukaihen [2003, s. 15]).

Loucopoulos [2000] jakaa tietojärjestelmien muutoksenhallinnan kahteen lähestymistapaan: jäsentävään eli *osittavaan* (engl. top-down) ja *kokoavaan* (engl. bottom-up). Osittavalla lähestymistavalla muutosprosessi alkaa korkean tason liiketoimintatavoitteiden määrittelyllä. Tämän jälkeen kehitysprosessin aikana vaatimuksia konkretisoidaan. Kokoavalla lähestymistavalla muutostenhallinta alkaa organisaation nykytilan analysoinnilla, todellisten prosessien analysoinnilla ja

määrittämällä, täyttävätkö olemassa olevat liiketoimintarakenteet sidosryhmien kiinnostuksen kohteet. Optimaalinen lähestymistapa olisikin näiden kahden lähestymistavan yhdistäminen. [Loucopoulos, 2000]

Myös tietovaraston kehittämismetodit voidaan jakaa osittavaan ja kokoavaan lähestymistapaan [Ponniah, 2001, s. 26-27]. Ponniah [2001, s. 26-27] toteaa, että osittavan lähestymistavan tavoitteena on kokonaisratkaisu, yrityksen laajuisen tietovaraston muodostaminen. Menetelmä pyrkii takaamaan palasten yhteensopivuuden korkeammalla tasolla asetettujen toiminnallisuusvaatimusten avulla. Päinvastainen kokoava lähestymistapa lähtee Ponniahin [2001, s. 26-27] mukaan yksikkökohtaisten paikallisvarastojen muodostamisesta.

Osittavan lähestymistavan käyttö on yrityksen yhteinen hanke. Osittavalla lähestymistavalla saadaan selville yrityksen näkemys hallitsemistaan tiedoista. Arkkitehtuurista saadaan siten itsenäinen eikä erilaisten paikallisvarastojen yhdiste. Tällöin myös metatiedoille on yksi ja keskitetty säilytyspaikka. Samoin tietovaraston säännöt ja hallinta ovat keskitettyjä. Iteraatioiden avulla toteutettuna voidaan myös saada aikaan nopeita tuloksia, mutta tietovaraston perustaminen vie tällöin myös kauemmin aikaa. Koska kyseessä on koko organisaation hanke, on mukana paljon ihmisiä ja tehtävä työ on laaja-alainen, ovat seurauksena myös korkeat kustannukset ja riskit ilman prototyyppejä ja varmistusta sovelluksen toimintakyvystä. Lisäksi toteutukseen tarvitaan yleensä korkean tason moniosaajia, kuten ihmisiä, joilla on liiketoimintaosaamista, raportointikokemusta sekä tietokanta-asiantuntemusta. [Ponniah, 2001, s. 26]

Kokoava lähestymistapa keskittyy itsenäisten paikallisvarastojen perustamiseen. Paikallisvarastot perustetaan inkrementaalisesti ja niiden perustamisjärjestys määrittäytyy yrityksen määrittämän prioriteettijärjestyksen mukaan, jolloin tärkeimmät paikallisvarastot voidaan toteuttaa ensimmäiseksi. Paikallisvarastot ovat pienempiä ja hallittavampia palasia, joten niiden toteuttaminen on helpompaa ja nopeampaa, jolloin epäonnistumisen riskikin on pienempi. Tällöin myös rahallisesta sijoituksesta ja prototyyppisovelluksesta tulee edullisempi, koska suunniteltava ja toteutettava kokonaisuus on pienempi verrattuna osittavaan lähestymistapaan. Pienempien toteutusjaksojen ansiosta projektiryhmän on mahdollista oppia ja kasvaa, jolloin opittua voidaan hyödyntää seuraavassa askeleessa. [Ponniah, 2001, s. 27]

Suurimpana haittana kokoavan lähestymistavan käytölle Ponniah [2001, s. 27] näkee tiedon pirstoutumisen, koska yksittäiset paikallisvarastot eivät yleensä ota huomioon koko organisaation vaatimuksia, joten niiden näkemys tiedosta on kapea-alainen. Hän toteaa, että kokonaisvaltaisesti organisaation tasolta katsottuna yhteensovittamaton sekä yhteensopimaton data lisääntyy ja toisteinen data leviää jokaiseen paikallisvarastoon. Myös hallitsemattomien rajapintojen määrä kasvaa, koska jokainen paikallisvarasto huolehtii omasta tiedonlatauksesta.

Ponniah [2001, s. 27] esittelee käytännöllisen lähestymistavan tietovaraston perustamiseen. Tämä lähestymistapa huomioi sekä osittavan, että kokoavan lähestymistavan hyvät puolet ja lähtee organisaation pitkän tähtäimen tarpeista. Nämä vaiheet on jaettu seuraavasti:

- 1.) Suunnitellaan ja määritetään tietovaraston vaatimukset yleisellä yrityksen laajuisella tasolla.
- 2.) Täydelliselle tietovarastolle luodaan ympäröivä arkkitehtuuri.
- 3.) Tietovaraston tietosisältö tehdään yhdenmukaiseksi ja yhtenäiseksi.
- 4.) Tietovarasto toteutetaan sarjana paikallisvarastoja, yksi kerrallaan.

Tietovaraston vaatimusten määrittämiseen Paim ja muut [2002] ehdottavat seuraavia tekniikoita: haastattelut, prototyypitestaukset ja skenaariot. Päätösten tekijöiltä kysytään kysymyksiä strategiseen analysointiin liittyen. Prototyypeillä esitetään päätöksentekijöille, kuinka järjestelmän toiminnot auttavat päätöksenteossa. Skenaariot, jotka toteutetaan käyttötapauksien, eli tietovarastosta raportoitavien liiketoimintatietojen kuvaamisen avulla, auttavat kehittäjiä selventämään ja tarkentamaan järjestelmän vaatimuksia. Koska tietovarastojärjestelmän perustoiminnallisuudet (kuten tietojen lataus, integrointi ja haut) toimivat projektista toiseen samalla tavoin, käyttötapojen kuvaamisella keskitytään toiminnallisuuteen, joka eroaa muista projekteista. [Paim *et al.*, 2002]

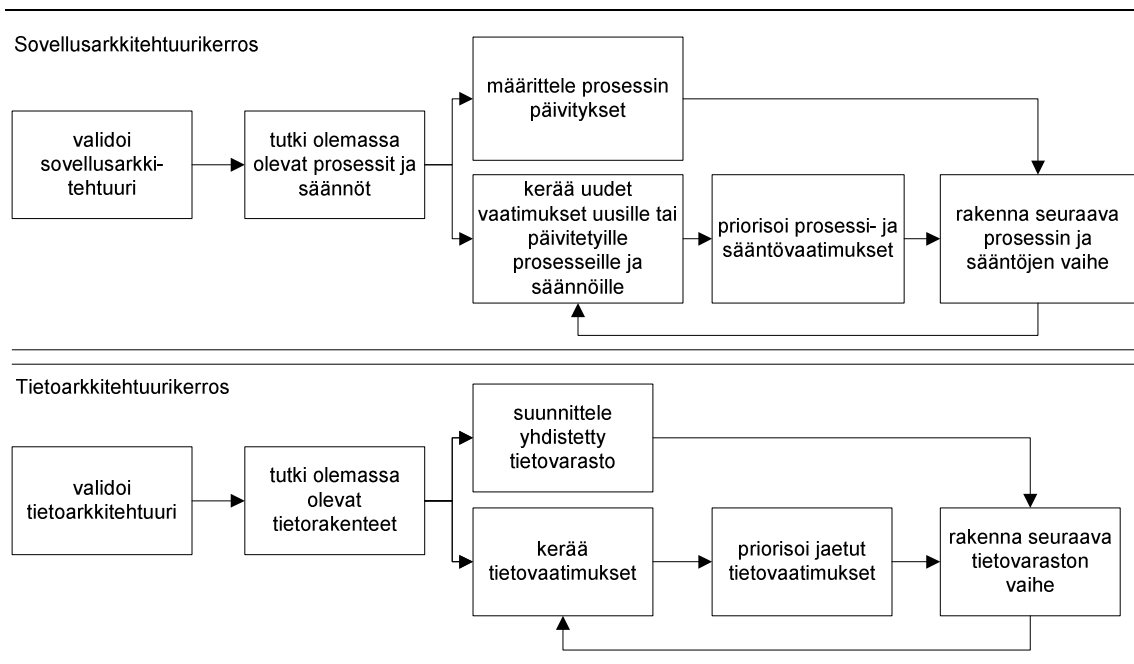
Tietojen latausvaihe on vaativin ja aikaavievin kehitysprosessin osuus [Shin, 2002]. Automaattisen ETL-työkalun käyttäminen on tietovaraston rakentamisen elinehto. Latausvaiheen toteuttaminen manuaalisesti ohjelmoimalla aiheuttaa päällekkäisyyttä, epäjohdonmukaisuutta sekä väärinmuotoiltuja ja laittomia arvoja sekä paljon ylimääräistä työtä [Shin, 2002]. Huolimatta tiedonpoiminnan ongelmista, tietovaraston päivittämisestä aiheutuvat kertautuvat kulut muodostavat tietovaraston kokonaiskustannuksista jopa 55 prosenttia [Vassiliadis, 2000].

Tietojen puhdistaminen on tärkeää, koska tietoja käytetään päätöksentekoon, jolloin virheelliset tiedot voivat johtaa virheellisiin tulkintoihin ja edelleen virheellisiin päätöksiin. Tietojen puhdistukseen tarkoitettut työkalut auttavat tunnistamaan tietojen poikkeamia ja helpottavat niiden korjaamista. Esimerkiksi käsitteellisessä tietomallissa valinnaisiksi kuvatut tiedot ja näistä mahdollisesti johdetut loogisen tietomallin attribuutit ovat yksi merkittävimmistä epäyhtenäistä tietoa sisältäviä paikkoja. [Chaudhuri and Dayal, 1997]

Tiedon keruun, puhdistamisen ja muunnosten jälkeen data ladataan ja siirretään tietovarastoon [Chaudhuri and Dayal, 1997]. Chaudhuri ja Dayal toteavat, että tyyppillisesti lataus suoritetaan eräajoissa, öisin ja pienessä aikaikkunassa, mutta suurten tietomassojen takia tiedonsiirto voi kestää hyvinkin kauan. He lisäävät, että on myös päätettävä kuinka usein ja milloin tietovarastoa päivitetään, koska riippuu tietojen

ajantasaisuus. ETL-vaiheessa yksi kriittinen tekijä projektin onnistumisen kannalta on tiedon keruun ja muunnostöiden ajallistaminen. Lisäksi Trisolini ja muut [1999] huomauttavat, että 10–15% muunnostehtävistä vaatii erityiskäsittelyä.

Shin [2002] esittelee eräässä projektissa käytettyä toteutusstrategiaa ja sen perusteella tehtyjä johtopäätöksiä. Projektien alussa on toteutettava tiukkaa teknisten, tieto-, sovellus- ja organisaatioarkkitehtuurien validointia. Organisaatioarkkitehtuurin validoinnissa on tarkasteltava kokonaisinformaationhallintastrategiaa, joka koostuu tiedonhallinnasta, rooleista, vastuista, liiketoiminta-alueiden informaatiovaatimuksista ja tarpeellisista organisaatiomuutoksista. Teknisen arkkitehtuurin evaluomisessa havaittiin tarpeelliseksi keskittyä OLAP-sovelluksen päivittämiseen, metatietojen hallintaan ja ETL-työkalujen käyttämiseen. Tieto- ja sovellusarkkitehtuurien vaihetuotteet perustuvat olemassa olevien tietorakenteiden ja sovellusvaatimusten iteratiiviseen arviointiin ja priorisointiin. Tieto- ja sovellusarkkitehtuurien arviointiaskeleet ovat esitetty kuvassa 21. [Shin, 2002]



Kuva 21. Tieto- ja sovellusarkkitehtuurien iteratiivinen validointi esitettyinä tietovirtakaavion avulla.

Ennen kuin tietovaraston toteuttaminen aloitetaan, on olennaista, että toteutuksen hallinnoimiseen on käytössä siihen soveltuva arkkitehtuuri ja menetelmä [Trisolini *et al.*, 1999].

4.2.1. Tietovarastointiprosessin sidosryhmät

Tietomallin valinnan vaikeus johtuu eri ihmisten erilaisista tarkastelukulmista, toteavat Moody ja Shanks [1994]. Esimerkiksi sovelluskehittäjä arvostaa heidän mukaansa toteutuksen helppoutta ja käyttäjä vaatimusten täyttymistä. Kummankin sidosryhmän näkemykset ovat oikeita, mutta ne eivät yleensä ole yhdenmukaiset. Jotta järjestelmästä

saadaan tehokas, on kaikkien sidosryhmien osallistuttava tietovaraston suunnitteluprosessiin ja heidät on myös saatava tyytyväisiksi [Moody and Shanks, 1994].

Tietovarastointiprojektin eri vaiheissa on mukana ihmisiä, jotka rooliensa perusteella voidaan jakaa esimerkiksi seuraavasti: päätöksentekijä, tietovaraston ylläpitäjä, tietovaraston suunnittelija ja tietovaraston ohjelmoija [Jarke *et al.*, 1998]. Inmon ja muut [1998, s. 316] jakavat projektiryhmän roolit ja vastuut viiteen kategoriaan: tietokantasuunnittelijoihin, tiedon arkeologeihin, järjestelmäohjelmoijiin, käyttäjiin ja kouluttajiin sekä tietovaraston ylläpitäjiin. Moody ja Shanks [1994] tunnistavat sidosryhmiksi liiketoimintaihmiset, tietojen analysoijat, sovelluskehittäjät ja tietosuunnittelijat.

Liiketoimintaihmiset tarvitaan tietovaraston luontiprosessiin mukaan siksi, että he edustavat omistajan ja asiakkaan näkemystä. Liiketoimintaihmiset ovatkin kiinnostuneita tietomallin laajuudesta; nykyisten ja tulevien tarpeiden täyttymisestä. Tietojen analysoijat edustavat puolestaan arkkitehdin näkemystä. Heidän vastuullaan on tietomallin kehittäminen ja eri sidosryhmien vaatimusten täyttäminen yhdessä hyvien käytäntöjen kanssa. Sovelluskehittäjät vastaavat tietomallin toteuttamisesta. Tietomallia he arvioivat projektivaatimustensa perusteella: onnistuuko toteuttaminen ajan, budjetin, resurssien ja teknologian puitteissa. Tietosuunnittelijat ovat vastuussa tietomallin integroimisesta yrityksen tietomalliin ja jakamisesta koko yritykseen. [Moody and Shanks, 1994]

Varsinkin suuremman yrityksen tietovarastointiprojektissa on huomioitava, että mahdollisesti kaikki organisaation tietotekniikkakumppanit ovat jollakin tavalla osallisina tietovarastointiprojektissa ja tällöin projektin monimutkaisuus kasvaa [Trisolini *et al.*, 1999].

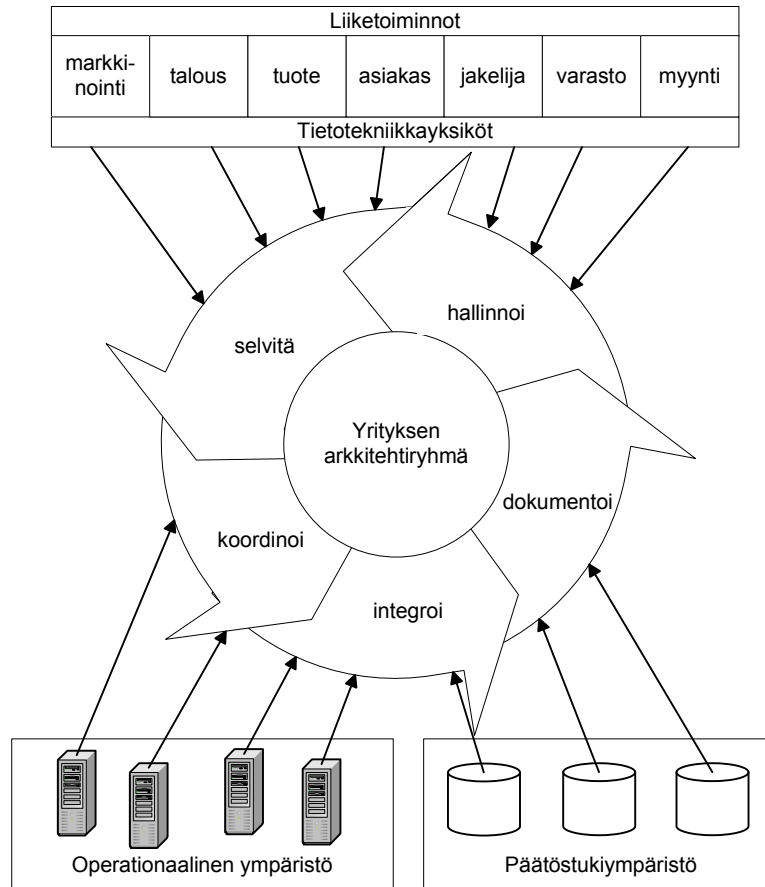
4.2.2. Organisaatiolähtöinen kehittämisprosessi

Tietovarastoinnin haasteet organisaation kannalta ovat entistä enemmän organisaatiolähtöisiä [Herrmann, 2004]. Yleensä tekniset asiat saavat helpommin huomion ja pääpainon pehmeämpien asioiden kustannuksella. Monet tutkimukset ovat kuitenkin osoittaneet, että organisatoriset, poliittiset ja kulttuuriset seikat ovat vähintään yhtä tärkeitä kuin tekniset asiat [Herrmann, 2004]. Herrmann on osoittanut tutkimuksessaan, että mitä pitkäkestoisempi hanke ja kypsempi tietovarasto sitä enemmän haasteet painottuvat organisatorisiin asioihin. Erilaiset organisatoriset ongelmat, kuten prosessien ja toiminnan puutteellisuus tai puuttuminen kosketti yli 60 prosenttia kyselytutkimukseen vastanneista [Herrmann, 2004]. Vain viidesosalla toiminta oli tarkasti kuvattu ja vain kolmasosalla oli prosessimallit kunnossa.

Moss ja Atre [2003, s. 67] suosittelevat, että yritykseen perustetaan arkkitehtuuriryhmä tai tiedonhallintaryhmä hallinnoimaan ja koordinoimaan yrityksen tietomallin toimintoja liiketoimintatiedonhallinta ja päätöksentekoympäristössä (kuva

22). Tällainen toimintamalli ottaa mielestäni paremmin liiketoiminnan vaatimukset huomioon ja liiketoimintaihmiset pääsevät paremmin osallistumaan tietovaraston kehittämisprosessin. Jos näin ei tehdä, tietotekninen toimintaympäristö sirpaloituu, kuten operatiivisille järjestelmille on käynyt. Ryhmän toimintaan kuuluu laajan liiketoiminta-analyysiin toteuttaminen yhdessä liiketoimintaihmisten kanssa. Jotta liiketoimintaihmisten osallistuminen olisi mahdollisimman kattavaa läpi organisaation, voidaan määrittelyvaiheessa soveltaa vertaisarviointia. Säännöllisen keskustelufoorumin avulla ja yhteistyössä liiketoimintaihmisten kanssa ylläpidetään ja tarkastetaan standardit ja liiketoimintasäännöt. Työn lopputuloksena organisaatiolle luodaan yksi tiivistetty ja toisteeton tietoarkkitehtuuri eli yrityksen looginen tietomalli. Se on pääväline operationaalisten tietojen kytkemiseksi tietovaraston tietoihin. [Moss and Atre, 2003, s. 67]

Menettelytavat, työjärjestykset, ohjeet ja standardit tarvitaan liiketoimintatiedonhallinta- ja päätöstukijärjestelmien ympäristön koordinointiin ja hallintaan. Ryhmä määrittää liiketoimintojen ja tietojen väliset suhteet ja samalla ratkaistaan mahdolliset olemassa olevat kiistat tietojen määrityksistä ja niiden arvoalueista eli validista sisällöstä. Tietojen nimet ja arvot standardoidaan, jotta ne vastaavat todellisia liiketoimintasääntöjä ja menettelytapoja. Oleellista on, että ryhmä ja liiketoimintaihmiset pääsevät yhteisymmärrykseen säännöistä ja menettelytavoista. Moss ja Atre [2003, 71, 76] toteavat, että liiketoimintatiedonhallintaympäristöjen on noudatettava arkkitehtuurisia standardeja samoin kuin esimerkiksi yrityksen www-sivuilla on yhtenäinen ulkonäkö. [Moss and Atre, 2003, s. 67]



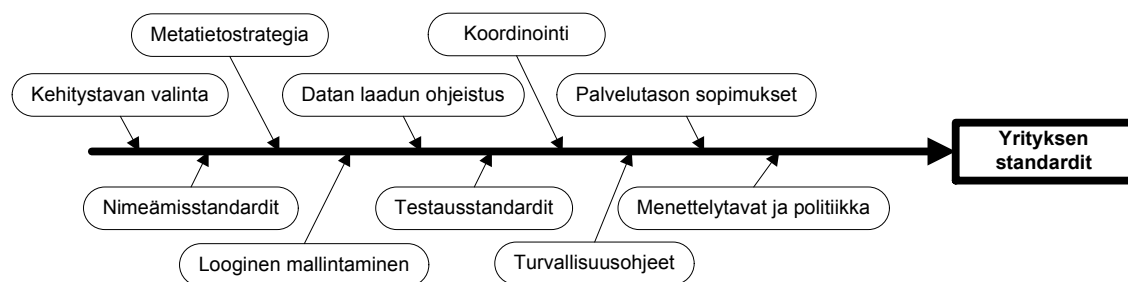
Kuva 22. Yritysarkkitehtuuriryhmän tehtävät ja sijainti organisaatiossa.

Ryhmän toimintaan kuuluu myös projektia tukeva ei-tekninen toiminta. Metatiedosta luodaan kuvauskanta ja siihen ladataan toisteetonta metatietoa. Lähtötiedot luetteloidaan ja kytketään soveltuviin liiketoimintatietokantoihin. Muut järjestelmän komponentit, kuten raportit ja ohjelmat, listataan, jotta voidaan tunnistaa voidaanko näitä komponentteja käyttää myöhemmin uudelleen. [Moss and Atre, 2003, s. 67]

Moss ja Atre [2003, s. 75] esittävät, että ei-teknisen infrastruktuurin toimenpiteet on toteutettava lineaarisesti. Projektien aluksi on arvioitava muun kuin teknisen infrastruktuurin komponenttien tehokkuus. Tämän jälkeen infrastruktuuri on dokumentoitava. Dokumentoinnin jälkeen infrastruktuuria voidaan tarvittavilta osin parantaa.

Infrastruktuurin arviointiraporttiin kuvataan nykytilanne ja puutteet sekä annetaan parannusehdotuksia. Raportissa voidaan lisäksi kuvata kehitysmetodin käyttötapa, arviointiohjeet, projektin laajuus, tehtävienhallinnan menetelmät, roolit ja vastuut, turvallisuusprosessi, metatiedonkeräys ja jakaminen, projektikohtaisen loogisen tietomallin sulauttamisprosessi yrityksen tietomalliin, tiedon laadun mittarit, testausprosessi, palvelutason sopimukset, tukifunktiot, kiistojen ratkaisumenetelmä sekä kommunikointiprosessi. Jos projektissa ei ole huomioitu kaikkia tarvittavia asioita, ne

ovat priorisoitava seuraavaan projektiin. Projektisuunnitelmaan on myös muistettava varata tarvittava aika parannusehdotuksien toteuttamiseksi ja puutteiden korjaamiseksi. [Moss and Atre, 2003, s. 76]



Kuva 23. Yrityksen ei-teknisen arkkitehtuurin evaluoinnin askeleet kuvattuna syy-seuraus-kaaviona.

Liiketoimintatiedonhallintaprojektien olosuhteet, laajuus ja toteutus vaihtelevat projekteittain, joten projekteissa pitää hyödyntää **soveltuva toteutustapaa**. Kuitenkin, jotta projektien riskejä voidaan hallita, on olemassa oltava ohjeet, joissa käydään läpi pakolliset vähimmäistoimenpiteet ja toimitettavat tuotokset, projektin päättymisvaatimukset sekä prosessin riippuvuudet. [Moss and Atre, 2003, s. 72]

Tietojen standardimainen nimeäminen ja lyhennestandardit tarjoavat yhtenäisen ja yhteisen liiketoimintatiedonhallintasovelluksien käyttötuntuman sekä kehittäjille, että liiketoimintaihmisille. Nimeämisen pitää olla tarkoituksenmukaisia ja liiketoimintaihmisille soveltuvia. Organisaatio voi luoda itselleen omat standardit tai soveltaa koeteltuja standardeja tai teollisuusstandardien nimeämiskäytäntöjä. Yrityksen laajuiset ja teollisuudelle ja organisaatiolle ominaiset lyhenteet ovat fyysisiä nimiä (kuten sarakkeiden ja taulujen nimet), ja ne tiedotettava koko organisaatiolle. [Moss and Atre, 2003, s. 72, 76]

Organisaatio tarvitsee metatietostrategian; standardit tai ohjeet, jotta voidaan vasta kysymyksiin, kuinka, kuka, milloin ja mitkä **metatiedon komponentit** ovat kerätty. Metatietoja ovat kuvailevat tiedot mm. liiketoimintaprosesseista, tietoalkioista, liiketoiminnan säännöt ja tiedon laatu. Metatiedon kuvauskannan on myös tuettava metatiedon keräämis- ja käyttöstandardeja. [Moss and Atre, 2003, s. 73, 76]

Projektikohtaiset tietomallit on yhdistettävä yhdeksi yhtenäiseksi ja integroiduksi yrityksen **loogiseksi tietomalliksi**. Organisaation on luotava lisäksi standardit projektikohtaisten loogisten mallien luomiseksi ja yhdistämiseksi yrityksen loogiseen tietomalliin. On myös varmistettava, että työryhmässä on riittävästi osallistujia [Moss and Atre, 2003, s. 73, 76]

Tietovaraston **tietojen laatu** voi olla korkeintaan yhtä hyvää kuin jalostamattomien lähtötietojen, joihin se perustuu. Useimmilla organisaatioilla on enemmän virheellistä ja epäpuhdasta tietoa (eli esimerkiksi puutteellisesti täytettyjä tietoja), kuin sitä voidaan puhdistaa. Organisaatioiden on suositeltavaa ohjeistaa, luokitella ja priorisoida tiedon

puhdistaminen ja jalostaminen. Hyväksyttävät laaturajat ja tiedonlatauksen eli ETL-toiminnallisuuden laadun mittaus on myös määritettävä. Koska tietovarastossa kuitenkin törmätään virheellisiin tietoihin, on standardisoitava myös virheidenkäsittely sekä virheellisten ja puuttuvien tietojen siirto myöhemmin käsiteltäväksi. [Moss and Atre, 2003, s. 73]

Testausstandardit puolestaan määrittävät minkä tyyppisiä testejä projektin aikana suoritetaan ja ketkä niihin osallistuvat. Testausta auttavat testipohjat ja ohjeet testitapausten määrästä ja kattavuuden arvioinnista, testaussuunnitelmista sekä kuinka eri testitoiminnot organisoidaan ja hallitaan. [Moss and Atre, 2003, s. 73, 76]

Liiketoimintatiedon hallintaympäristössä on useita kohdetietokantoja ja useita sovelluksia. Koska liiketoimintatiedonhallintasovellukset eivät ole itsenäisiä järjestelmiä, niiden kehitystä on **koordinoitava** ja yhteen sovitettava, jotta voidaan taata päätöksentekoympäristön eheys. Se tarkoittaa esimerkiksi yhtä loogista keskuslatausaluetta ja moduulia ETL-prosesseille ei paikallisvarastokohtaisia. [Moss and Atre, 2003, s. 74]

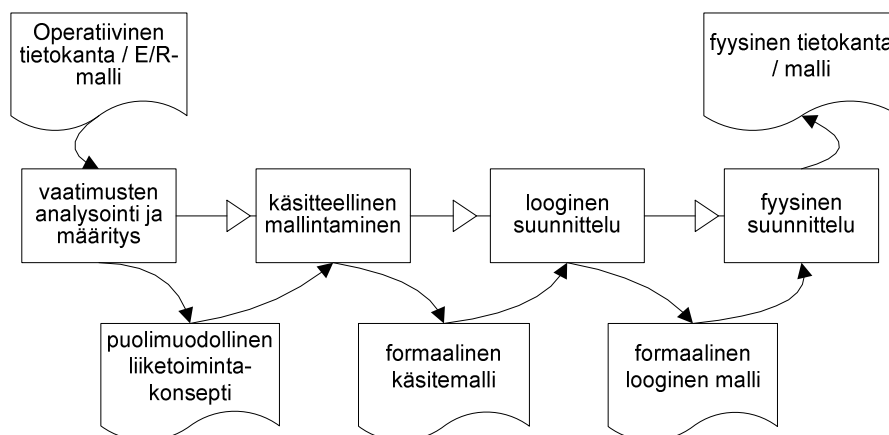
Tietovaraston tieto on peräisin operationaalisista järjestelmistä, siksi **turvallisuusohjeiden**, jotka koskevat operationaalisia tietoja, pitäisi koskea myös liiketoimintatiedonhallintaympäristön tietoja. Tietovaraston mitattavat tiedot ovat yleensä kuitenkin summattu, joten yksityiskohtia ei ole niin paljon näkyvissä ja turvallisuustasoa voi laskea. Tietojen omistajien on kuitenkin luotava turvallisuusohjeet määrittämään pakolliset turvallisuustoimenpiteet, jotta tietojen luvaton ja vahingossa tapahtuvaa katselua voidaan hallita. Turvallisuusriskit voidaan luokitella esimerkiksi tietojen arkuuden ja ohjelmistojen käyttötarkoituksen perusteella. [Moss and Atre, 2003, s. 74]

Palvelutason sopimuksilla (engl. service-level agreement, SLA) määritellään liiketoimintaperiaatteet ja niiden minimihyväksyttävyytasot. Organisaatiot toimivat eksplisiittisten ja implisiittisten liiketoimintaperiaatteiden mukaan. Esimerkiksi itseoppinut henkilökunta on implisiittinen periaate, kun eksplisiittinen puolestaan näkyy esimerkiksi yrityksen julkilausutussa visiossa. Kaikkien projektien on saavutettava talousdatan 98 prosentin laatutaso” on esimerkki palvelutasosopimuksesta [Moss and Atre, 2003, s. 74]

Standardien ja ohjeiden on katettava myös organisaation **menettelytavat ja politiikka**, kuten muutoksienhallintaproseduurit ja tehtävienseuranta- ja hallintatoiminnot. Myös kommunikointitavat, arviointiohjeet, roolit ja vastuut sekä dokumenttimallipohjat ovat osa menettelytapoja. Poliitiikan, menettelytapojen sekä standardien ja ohjeiden tarkoituksena on virtaviivaistaa sekä standardoida liiketoimintatiedonhallinnan sekä päätöstukijärjestelmien järjestelmät. [Moss and Atre, 2003, s. 75]

4.2.3. Suunnittelumallit

Hüsemann ja muut [2000] ehdottavat tietovaraston suunnitteluprosessiksi muokattua versiota perinteisestä tietokannan suunnitteluprosessista. Se koostuu vaatimusten analysoinnista ja määrittämisestä, käsitteellisestä mallintamisesta, loogisesta suunnittelusta sekä fyysisestä suunnittelusta (kuva 24).



Kuva 24. Eräs tietovaraston suunnitteluprosessi vuokaaviolla kuvattuna.

Kuvaan kahta ensimmäistä vaihetta tarkemmin, koska keskityn tutkimuksessani prosessin ylempiin tasoihin: vaatimusmäärittelyyn ja käsitteelliseen mallintamiseen. Operatiivisen E/R-mallin avulla liiketoiminta-alueiden asiantuntijat valitsevat oleelliset operatiivisen tietokannan attribuutit ja määrittävät niille tarkoituksen. Jokaiselle attribuutille määritetään valinnaisuus ja tarkoitus: joko dimensio tai mitta. Samalla lisätään täydentävät vaatimukset edelleen johdettavista mittareista ja liiketoiminnasta aiheutuvat rajoitteet sekä mahdolliset dimensioattribuuttien ominaisuudet [Hüsemann *et al.*, 2000].

Hüsemann ja muut [2000] määrittävät käsitteellisen tietovarastomallintamisen tarkoituksiksi käsittekaavion suunnittelun ts. tuottaa graafinen moniulotteinen kaavio liiketoimintavaatimusmäärittämisestä. He ovat luoneet mallintamisesta prosessimallin ja jakavat sen kolmeen peräkkäiseen vaiheeseen:

- 1.) asiayhteyksien määrittäminen mitattaville tekijöille (faktoille)
- 2.) dimensiohierarkioiden suunnitteleminen
- 3.) summattavien rajoitteiden määrittäminen.

Kaavion on tarkoitus yhdistää faktakaavio dimensiohierarkioihin ja mitattaviin asioihin. Summattavien rajoitteiden mitattavuus voidaan kuvata esimerkiksi taulukkona tai tietokuutiona [Hüsemann *et al.*, 2000].

4.2.4. Tietovaraston jatkuva kehittäminen

Tietovarasto on harvoin staattinen tietojärjestelmä, vaikka operationaaliset tietojärjestelmät ovatkin yleensä staattisia [Inmon *et al.*, 1997, s. 326]. Tietovarasto on

myös erittäin monimutkainen järjestelmä, jonka komponentit elävät usein ja riippumatta toisistaan [Vassiliadis *et al.*, 2000]. Tietovarasto vaatii jatkuvaa ylläpitoa ja tietovaraston on reagoitava ympäristön ja yrityksen liiketoiminnan muutoksiin. Liiketoiminnan muutoksia ovat esimerkiksi fuusioituminen, uuden liiketoimintayksikön perustaminen tai liiketoimintayksikön myynti toiselle yritykselle sekä lakimuutokset. Ympäristön muutoksia ovat esimerkiksi uuden järjestelmän käyttöönotto tai järjestelmän käyttämän tietokannan päivitys. Tekninen ympäristö muuttuu silloin, kun tuotteita kehitetään, päivitetään tai tuotevalikoimaa muutetaan [Vassiliadis *et al.*, 2000]. Muutoksia aiheuttavat myös lisääntynyt tiedon tarve ja käyttäjät yleensä pyytävät — täydellisestä tietovarastosta huolimatta — lisää tietoja [Inmon *et al.*, 1997, s. 326]. Vassiliadis ja muut [2000] muistuttavat, että käyttäjien vaatimukset muuttuvat alinomaa, joten oletetut raja-arvot ja priorisoinnit muuttuvat ajan myötä. He toteavat, että käyttäjät voivat myös luoda uusia ja päivittää vanhoja näkymiä. Toisaalta tiedonlähteitä voi myös kadota samalla, kun uusia lisätään järjestelmään.

Organisaation kattavaa mallia kehitetään, kun yrityksen tavoitteet tai strategia vaihtuvat, toisin sanoen organisaation liiketoimintasäännöt eivät pysy samoina, koska reaali maailma muuttuu [Vassiliadis *et al.*, 2000]. Inmon ja muut [1997, s. 326] muistuttavat, että mahdollinen katastrofi on vältettävissä vain suunnittelemalla tietovaraston muutostahti etukäteen. Tietovaraston käyttötarve ei yleensä lopu koskaan ja vaikka tietovaraston käyttö joskus loppuisikin, sen sisältämä data todennäköisesti säilyy ja vain sen säilytyspaikka vaihtuu. Säilytyspaikan vaihtuessa tiedoille saatetaan kuitenkin suorittaa uudelleenorganisointi.

Tietovaraston muuttuvassa ympäristössä yrityksen tavoitteiden ja strategian arvioinnit ovat riippuvaisia toisistaan [Vassiliadis *et al.*, 2000]. Vassiliadis ja muut [2000] näkevät siihen kolme pääsyötä:

- tietovaraston jatkuva kehittäminen: luonnollisia muutoksia tapahtuu monimutkaisessa ympäristössä
- laatuavoitteissa epäonnistuminen
- metalaatu: mittausprosessin laadusta ei ole varmuutta.

Täydentävä metatieto on ratkaisu tietovaraston kehittämisen kontrollointiin. Metatiedon talletuspaikka kerää muutoshistorian ja metatiedon avulla saadaan yhdenmukaiset säännöt laatutekijöiden uudelleenarviointiin. Kerätyn metatiedon avulla voidaan arvioida tietovaraston kehitykseen vaikuttavien tekijöiden ja toimenpiteiden vaikutusta muihin tekijöihin ja niiden seurauksista laatuun.[Vassiliadis *et al.*, 2000]

Liiketoimintasäännöt voivat muuttua jopa tietovaraston rakentamisen aikana. Tietovaraston jatkuva muuttuminen aiheuttaa monenlaisia ongelmia. Projektin kustannukset kasvavat ja aikataulu venyy. Kehitysryhmä turhautuu. Metatiedonhallinnan puuttuminen tekee melkein mahdottomaksi tunnistaa tietovaraston

kohdat, jotka pitää päivittää. Metatietojen puuttuminen ja ison muutoksen tekeminen saatiin eräässä projektissa toteutettua vain käytössä olleella yhdenmukaisella nimeämisellä. [Vassiliadis, 2000]

4.2.5. Metatieto ja tietovarastointi

Metatieto on tietoa tiedosta [Watson and Gray, 1998, s. 15]. Lamb [2001] jakaa metatiedon kahteen keskeiseen käyttötapaan: käyttäjien ja tietokoneiden tarpeisiin. Metatiedon päämääränä on:

- 1.) tarjota informaatiota, jota ihmiset kykenevät lukemaan; joka helpottaa elektronisen tiedon löytämistä ja tulkintaa monimutkaisessa ympäristössä.
- 2.) tarjota tietokoneiden prosessoitavissa olevaa dataa, joka helpottaa tietojen vaihtoa järjestelmien välillä ja tietojen prosessointia järjestelmissä.

Tietovarastoinnin näkökulmasta metatieto kattaa kuvaustiedon tietosisällöstä ja tietorakenteista sekä tietojen prosessointiin liittyvän informaation [Staudt *et al.*, 1999a]. Metatiedon tuottaminen ja hallinta vähentää tietovaraston kehityksen ja hallinnon kuormaa sekä parantaa tietovarastosta saatavan tiedon laatua ja määrää [Staudt *et al.*, 1999b].

Metatiedon rooli tietovarastointiprosesseissa kattaa järjestelmän kaikki elinvaiheet. Metatietoa voidaan luokitella sen käyttötavan mukaan. Passiivista metatietoa käytetään tietovaraston rakenteen, kehitysprosessin ja tietovaraston käytön yhtenäisessä dokumentoinnissa. Tämä dokumentointi kattaa kaikki käyttäjätasot ja niiden tehtävät. Puoliaktiivinen metatieto sisältää sekä aktiivisen, että passiivisen metatiedon piirteitä. Tällainen metatieto on staattisesti säilytettyä informaatiota, kuten rakennemääritelmiä ja kokoonpanomäärittelyksiä, joita luetaan ohjelmallisesti suoritusvaiheessa. Esimerkiksi kyselyjäsentäjien tarvitsemaa tietoa jonkin attribuutin olemassaolosta voidaan lukea ajoaikana. Aktiivinen metatieto on semanttista informaatiota, kuten tietojen muunnossääntöjä, joita tulkitaan ja suoritetaan ajonaikaisesti. Näitä tietoja säilötään keskitetysti kuvauskannassa ja hallitaan aktiivisesti. Tällaisen metatieto-ohjatun tietovarastointiprosessin dokumentaatio on yhtenäistä ja ajantasaista. [Staudt *et al.*, 1999b]

Metatiedoilla on tämän tutkimuksen kannalta keskeinen merkitys, koska niiden rooli tietovarastointiprosessin eri vaiheissa on keskeinen ja organisaatio-orientoituneen lähestymistavan hyödyntämisen edellytys. Prosessin osallisten tarvitsema dokumentaatio on pääosin passiivista metatietoa, mutta aktiivisenkaan metatiedon merkitystä ei voida sivuuttaa. Metatietoja voidaan käyttää myös uuden työntekijän perehdyttämiseen yritykseen tai tiettyyn liiketoimintasektoriin.

Staudt ja muut [Staudt *et al.*, 1999b] esittelevät metatiedon merkitystä kehitys- ja ylläpitovaiheelle. Järjestelmäintegraation mahdollistumiseen liittyvät keskeisesti tietolähteiden ja kohdejärjestelmien rakenteiden ja tarkoitusten tunteminen sekä

muunnossääntöjen luominen näiden järjestelmien tietojen välille. Metatieto lisää myös sovelluskehitysprosessin ohjattavuutta sekä luotettavuutta, koska tiedon merkitys, rakenne ja alkuperä ovat määritelty. Aktiivinen ja puoliaktiivinen metatieto auttavat ohjelmiston eri osien uudelleenkäytettävyyttä. Uudelleenkäytön avulla lisätoiminnallisuuden tekeminen on helpompaa. Ylläpitotehtävien, kuten tietojen latauksien, automatisoinnissa metatieto on keskeisessä roolissa ja samalla sitä voidaan tallentaa lokitarkoituksia varten. Koko tietovarastojärjestelmän turvallisuusmekanisminkin voidaan hallita metatiedon avulla.

Shin [2002] toteaa tutkimuksessaan, että metatiedonhallinta on erittäin tärkeä osa projektia. Vanhentunut metatieto ja metatiedon hallintastrategian puuttuminen aiheuttivat jatkuvasti ongelmia Shinin [2002] kuvaamassa projektissa. Metatietokeskeisen järjestelmän suunnittelu ja toteuttaminen metatietotyökalujen avulla pakottavat loogisen metatiedon eheyteen ja johdonmukaisuuteen [Shin, 2002]. Lisäksi työkalut tarjoavat dokumentaation tietojen muunnossäännöistä ja mahdollistavat tiedonmäärittämissä automaattisen generoimisen [Shin, 2002].

Metatiedon avulla saadaan selville, mitä tietovarastossa on. Metatieto on oleellinen työkalu sekä käyttäjille että tietovaraston tekijöillekin, vaikka he tarvitsevat erilaista tietoa. Operatiiviset tiedot liitetään tietovaraston muotoon metatiedon ohjeiden avulla. Ohjeiden avulla tiedot voidaan muuntaa tietovaraston käytäntöjä noudattaen oikeaan ja standardiin muotoon. Tiedot summataan metatiedon sääntöjen mukaan. Käyttäjille näkyviä metatietoja ovat liiketoimintatermit, joilla tietoa kuvataan sekä liiketoimintatermejä vastaavat tekniset nimet. Muita käyttäjää kiinnostavia metatietoja ovat tietojen alkuperä, tietojen johtamiseen käytetyt säännöt sekä tiedon luontiajankohta. [Watson and Gray, 1998, s. 15]

4.3. Tietovaraston kehitysmetodit

Tietovarastoprojektit eroavat perinteisistä ohjelmistokehitysprojekteista. Merkittävin ero on projektin laajuus: Adelman ja Moss [2000, s. 16] toteavat, että perinteiset projektit palvelevat vain tiettyä liiketoimintayksikköä, kun taas tietovarastoprojektit palvelevat koko organisaatiota. Adelman ja Moss [2000, s. 16] tarkentavat, että tietovarastoprojektien kehitys tapahtuu yli organisaatorajojen; tällöin toimintatapojen on muututtava ja totuttuihin kehitysmetodologioihin on tultava muutoksia.

4.3.1. Tietovaraston kehittämisen erityispiirteitä

Tietovaraston kehitysmetodin on oltava Kimball:n ja Ross:n [2002, s. 368] iteratiivinen (työvaiheiden toistaminen välillä arvioiden). He toteavat, että tietovaraston elinkaari on paljon pidempi kuin liiketoiminnan kulloisetkin toimintatavat. Lisäksi he pitävät organisaation tietopääomaa yleensä liian laajana ja monimutkaisena, jotta se voitaisiin määrittellä etukäteen, ennen tietovaraston luomista. Trisolini ja muut [1999] lisäävät, että tietovaraston luominen on inkrementaalinen prosessi, erityisesti tapauksessa, jossa

tietovaraston rakenne ja organisaatio ovat monikerroksisia. Inkrementaalinen lähestymistapa sekä tietointegraatioon että tietovarastointiin vaatii inkrementaalista lähestymistapaa tukevia suunnittelumetologioita ja työkaluja [Trisolini *et al.*, 1999].

Kimball ja Ross [2002, s. 368] jakavat tietovaraston kehittämisen kolmeen merkittävään osaan: liiketoimintavaatimusten keruuseen, toteutusvaiheeseen ja tietovaraston julkaisuun. He kuvaavat kehittäjien ja liiketoimintaihmissen vuorovaikutusta seuraavin tavoin: ensimmäisessä vaiheessa on liiketoimintaihmissen ehdotuksia kuunneltava tarkoin. Toisessa vaiheessa tapahtuvat muutokset aiheuttavat aikataulun venymisen. Kolmas vaihe jäädyttää toteutuksen. Kimball ja Ross [2002, s. 368] löytävät toiselle ja kolmannelle vaiheen muutoksille yhteisen piirteen: muutosten aiheuttama tietovaraston tarkoituksen hämärtyminen. Uusien toimintojen myötä työ laajenee ja alkuperäiset tavoitteet saattavat unohtua.

Tietovaraston kehittämisen onnistuminen riippuu käyttäjien vaatimusmäärittelyn tarkkuudesta, toteavat Paim ja muut [2002]. Tietovaraston informaation hyödyntäminen riippuu puolestaan tietojen onnistuneesta hyväksikäytöstä. Siksi tietojen analysointi on yksi oleellinen osa tietovaraston kehittämisprosessia. Analysoinnin avulla saadaan tiedot integroitua, kuvaavat Adelman ja Moss [2000, s. 234] analysoinnin merkitystä.

Käsite-orientoitunut tietojen integroiminen vaatii loppukäyttäjiltä ja tietovaraston kehitystiimiltä vahvaa tietämystä tietojen merkityksestä ja sisällöstä [Paim *et al.*, 2002]. Näistä seikoista johtuen Paim ja muut [2002] väittävät, että vaaditaan metodologinen lähestymistapa, jotta voidaan taata 1.) oikea käsitys ja laajuus käyttäjän vaatimuksista, 2.) hienosäädetty suunnitteluvaihe ja 3.) sellaisen dimensionaalisen mallin luominen, joka tekee mahdolliseksi päätöstentekijöiden tarpeelliset analysointitarpeet [Paim *et al.*, 2002].

Iteratiivista toteutustapaa käytettäessä on tietovaraston rakentamisen onnistuminen todennäköisempää, mikäli aluksi keskitytään Kimball:n ja Ross:n [2002, s. 368, 396] ehdotuksesta yhteen liiketoimintaprosessiin, yksinkertaisempiin ja itsenäisiin tietovaraston osiin: paikallisvarastoihin ja mahdollisesti jopa vain yhteen lähteeseen.

Tietovaraston luontiprosessin lisäksi myös tietojen integrointi on inkrementaalinen prosessi [Trisolini *et al.*, 1999]. Inkrementaalista eli asteittain täydentyvää lähestymistapaa Trisolini perustelee organisaatiotason näkemyksellä, koska kaikkia lähteitä ei ole mahdollista saada yhdistettyä kerralla. Trisolini korostaa, että inkrementaalisen luonteen takia on suunniteltava metodologiat ja työkalut, jotka tukevat tätä prosessia.

4.3.2. Lähestymistapoja tietovaraston kehittämiseen

Kaldeich ja Sá [2004] ovat ryhmitelleet neljä erilaista lähestymistapaa tietovaraston kehittämiseen ja ylläpitämiseen. Nämä kehitysmetodit ovat tieto-ohjattu (engl. Data-Driven), tavoiteohjattu (engl. Goal-Driven), vaatimusohjattu (engl. Demand-Driven) ja näiden yhdistelmä integroitu prosessiohjattu (engl. Integrated-Process-Driven)

lähestymistapa, joka on Kaldeichin ja Sán [2004] kehittämä organisaatioprosessilähtöinen tietovarastojärjestelmämetodologia. Näiden lähestymistapojen tarkoituksena on määrittellä tietovarastokäyttäjien informaatiovaatimukset ja vastata tulevaisuuden tietovaraston käyttäjien vaatimuksiin suoraan käytettävissä olevalla informaatiolla. Kaikkien metodien yhteisenä lähtökohtana on informaatio. Informaation tarjonta johtaa tietolähtöiseen lähestymistapaan, kun vastaavasti vaatimuslähtöisen lähestymistavan kehitysmetodin käyttäminen johtuu informaation tarpeesta. [Kaldeich and Sá, 2004]

Tietolähtöinen lähestymistapa perustuu organisaation tietomallien ja niihin liittyvien tapahtumien analysointiin. Tämä eroaa perinteisestä vaatimuslähtöisestä järjestelmäkehityksestä, koska tietovarastokehityksessä vaatimukset tunnistetaan vasta sen jälkeen, kun alkulatauksella on kerätty alustavat tietovaraston tiedot ja käyttäjät ovat analysoineet tietovaraston tietoja. Tietolähtöinen metodi ei myöskään huomioi organisaation tavoitteita ja käyttäjien vaatimuksia. Metodien ongelmana ovat käyttäjien motivointi tietomallien käyttämiseen ja tarpeettomien tietorakenteiden tutkimiseen haaskatut resurssit. [Kaldeich and Sá, 2004]

Vaatimuslähtöinen lähestymistapa olettaa, että organisaation tavoite on sama kaikille. Lähestymistavan tavoitteena on liiketoimintalähtöisen prototyypin muodostaminen. Liiketoimintaihmiset määrittelevät tavoitteet ja priorisoivat liiketoimintakysymykset, jotka tukevat näitä tavoitteita. Tämän jälkeen liiketoimintakysymysten avulla tunnistetaan tietovaraston tietoalkiot ja hierarkiat. [Kaldeich and Sá, 2004]

Tavoiteohjattu lähestymistapa on erityinen vaatimuslähtöinen lähestymistapa, jolla liiketoimintaprosessien yksityiskohtaisella analysoinnilla tavoitetaan informaation vaatimukset ja muunnetaan oleelliset liiketoimintaprosessien tietorakenteet tietovaraston tietorakenteiksi. Tavoiteohjatussa lähestymistavassa aluksi asetetaan tavoitteet ja määritetään organisaation palvelut. Sen jälkeen liiketoimintaprosessia peilataan asiakkaisiin sekä asiakkaan ja organisaation välisiin tapahtumiin. Tapahtumat muunnetaan seuraavaksi olemassa olevien tietojärjestelmien tapahtumiksi. Lopuksi tunnistetaan mitat ja dimensiot tietovarastoa varten. Yksityiskohtaisen liiketoimintaprosessien analysoiminen ei ole kuitenkaan tarkoituksenmukaista, koska käyttäjien tehtävät ovat yksilöllisiä ja vapaamuotoisia. Lisäksi päätöksentekijät usein kieltäytyvät paljastamasta prosessejaan yksityiskohtaisesti. [Kaldeich and Sá, 2004]

4.3.3. Tietovaraston kehitysmenetelmiä

Böhnlein ja Ende [2000] toteavat, että yleinen tapa tietovaraston moniulotteisten tietorakenteiden kehittämiseen on johtaa soveltuvat tiedot operatiivisista tietolähteistä. Böhnlein ja Ende [2000] näkevät, että tietovaraston rakentamisen yksi merkittävä ongelma on potentiaalisten käyttäjien informaatiotarpeiden tunnistaminen ja huomioon ottaminen. He lisäävät, että liiketoimintaprosessit ovat merkittävässä asemassa

yrityksen toiminnan ja organisaation analysoimisessa sekä suunnittelussa. Tietovaraston kehitystekniikoista voidaan tunnistaa esimerkiksi käyttäjäkeskeinen, operatiivisesti suuntautunut ja liiketoimintaprosessikeskeinen tietovaraston kehitysmenetelmä [Böhnlein and Ende, 2000].

Käyttäjäkeskeinen tietovarastointikehitys on yleensä kuvattu epämuodollisena tapana luoda tietovaraston rakenteita. Tämän kehitystavan merkittävimmät metriikat koostuvat organisaatiolle määritetyistä kriittisistä menestystekijöistä, jotka riippuvat yhtiön johdon strategiasta ja tavoitteista. Tietovaraston dimensiot riippuvat näiden menestystekijöiden vaikutuksista. Kehitystekniikassa ei kuitenkaan ole ennalta määrättyä tapaa näiden tekijöiden määrittämiseksi. Yksi tapa informaatiotarpeiden keruuseen ovat haastattelut. Kehitystavan haittatekijänä on usein käyttäjien tietämättömyys. Käyttäjät eivät yleensä osaa kuvata vaatimuksiaan täsmällisesti. Lisäksi useimmilla käyttäjillä on vain aavistus tarpeista ja useimmat käyttäjät eivät myöskään ole nähneet raportointityökaluja aikaisemmin. [Böhnlein and Ende, 2000]

Operatiivisesti suuntautunut tietovarastointikehitys on konkreettisempi tapa analysoida operatiivisen järjestelmän tietomalleja ja tunnistaa merkitykselliset transaktiot. Tämän kehitysmenetelmän ongelmana on tietomallien rajoittuneisuus: tietomallit kuvaavat yrityksen liiketoimintaa epätarkasti. [Böhnlein and Ende, 2000]

Liiketoimintaprosessikeskeinen tietovarastointikehitysmenetelmän keskeinen etu on sen avulla syntyvä liiketoimintaprosessimalli sisältää muodollisen kuvauksen käyttäjien informaatiovaatimuksista [Böhnlein and Ende, 2000]. Lisäksi hyötynä nähdään mahdollisuus tunnistaa informaatiovaatimukset, joita ei voida toteuttaa käytettävissä olevien lähtöjärjestelmien tietojen pohjalta [Böhnlein and Ende, 2000]. Esimerkiksi tiedot ovat lähtöjärjestelmässä liian epätarkkoja tai tietoja kerätään harvemmin kuin käyttäjien vaatimukset edellyttäisivät.

Böhnlein ja Ende [2000] esittelevät liiketoimintaprosessikeskeisen mallin tietovaraston alustavien rakenteiden muodostamiseksi. Se perustuu semanttiseen objektimalliin. Semanttinen objektimalli (engl. the Semantic Object Model, SOM) on perusteellinen ja integroitu metodologia liiketoimintasuunnitteluun. Se on kolmitasoinen liiketoiminnan arkkitehtuurinen kehys ja sitä voidaan käyttää analysointiin ja suunnitteluun. Yritysarkkitehtuuri on perusta SOM-metodologialle. Yritysarkkitehtuuri koostuu yrityksen suunnitelmasta, joka määrittää mallinnettavan järjestelmän ja tuotteet, palvelut sekä tavoitteet. Liiketoimintaprosessimalli on yritysarkkitehtuurin toinen taso. Kolmas taso on resurssien määrittäminen, jossa huomioidaan liiketoimintaprosessien tehtävät kuten henkilöstö, järjestelmät ja laitteet. [Böhnlein and Ende, 2000]

Böhnlein ja Enden [2000] liiketoimintaprosessikeskeinen lähestymistapa koostuu neljästä vaiheesta. Vaiheiden yhteinen piirre on palautteen muodostaminen aikaisemmille vaiheille.

- 1.) Järjestelmän palveluiden ja tavoitteiden määrittäminen. Vaiheen lopputuloksena ovat tavoitteet ja palvelut.
- 2.) Liiketoimintaprosessien analysointi. Vaiheen lopputuloksena ovat rajatut prosessit sovellusalueesta.
- 3.) Käsitteellisen objektimallin johtaminen, toisin sanoen liiketoimintaprosessimalli muunnetaan sovellustasolle. Vaiheen lopputuloksena tunnistetaan oleelliset ja analysoitavat attribuutit.
- 4.) Alustavien tietovarastorakenteiden tunnistaminen. Vaiheen lopputuloksena ovat mittarit, dimensiot ja rajoitteet. Nämä visualisoidaan tähti- tai lumihiutalemallilla.

Böhnlein ja Ende käyttävät tietovaraston alustavien rakenteiden määrittelyyn liiketoimintaprosessimallia operationaalisten tietomallien sijaan. Heidän esittelemä määrittelyketju - tavoitteet, palvelut, liiketoimintaprosessit, mitat - varmistaa, että järjestelmän tavoitteet täyttyvät ja liiketoimintaihmisille tärkeät mitat sekä niihin liittyvät dimensiot saadaan muodostettua. He muistuttavat, että tähtimalli perustuu relaatiotietokantajärjestelmiin, joten malli ei sellaisenaan sovellu tietovaraston rakenteiden kuvaamiseen käsitteellisellä tasolla. [Böhnlein and Ende, 2000]

Moody ja Kortink [2000] esittelevät yleisesti teollisuudessa käytetyn metodin tietovaraston ja paikallisvarastojen kehittämiseen yrityksen tietomallista.

- 1.) Kehitä yrityksen tietomalli, jos sellaista ei ole jo olemassa.
- 2.) Suunnittele mallin perusteella keskitetty tietovarasto (engl. Central Data Warehouse) valitsemalla tärkeimmät liiketoiminta-alueet päätöksenteon tueksi.
- 3.) Luokittele tietovaraston mallin käsitteet tapahtumiin, komponentteihin ja luokitteleviin. Tapahtumat kuvaavat tapauksen tietyssä ajan hetkenä, esimerkiksi tilaukset ja hotellivaraukset. Komponenttikäsitteet vastaavat kysymyksiin, kuka, mitä, milloin, missä, kuinka ja miksi ja niitä ovat esimerkiksi asiakas ja tuote. Luokittelevia ovat hierarkioiden osat, kuten asiakastyypit.
- 4.) Tunnista tietomallin hierarkiat, joissa käsitteet ovat liitoksissa toisiinsa yhden suhde moneen suhteilla ja ne ovat kohdistettu samaan suuntaan.
- 5.) Suunnittele paikallisvarastot käyttämällä tähtiklusterimalleja jokaiseen keskitetyn tietovaraston tapahtumakäsitteeseen.

Moodyn ja Kortinkin [2000] mukaan lähestymistavan hyötynä seuraa jäsennellympi lähestymistapa dimensionaaliseen mallintamiseen kuin käyttämällä ensimmäisiä peruskäsitteitä. Lähestymistapa myös varmistaa, että paikallisvarastot ja tietovarasto tukeutuvat operatiivisten järjestelmien oikeisiin tietojen suhteisiin. Lähestymistavalle oleellista organisaation tietomallia hyödyntämällä, poiminta- ja latausprosessit

yksinkertaistuvat, koska tällöin esimerkiksi tiedot haetaan vain kustakin tiedot omistavasta alkuperäisestä (engl. master) järjestelmästä. Olemassa oleva yrityksen tietomalli auttaa myös tunnistamaan informaatiovaatimukset kokoavalla tavalla, eli tunnistamalla yrityksen tietojen sijainnin. Yrityksen tietomalli on myös vakaampi perusta suunnitteluun kuin käyttäjien ailahtelevat ja arvaamattomat kyselyvaatimukset. Lähestymistapa myös varmistaa, että tietovarastosta muodostuu riittävän joustava analysointivaatimuksille, koska tieto tallennetaan yksittäisten tapahtumien tasolla. Menetelmä myös maksimoi tietojen eheyden eli tietojen liittymisen toisiinsa. ER-mallin käyttäminen parantaa dimensionaalisten mallien kehittämisen lähtökohtaa ja tarjoaa tukea kokemattomille suunnittelijoille. Lisäksi analysointivaiheen jälkeen dimensionaalisen mallin kehittäminen on melko suoraviivaista. [Moody and Kortink, 2000]

5. Tapaustutkimukset

Tämä tutkielma käsittelee kahta tapausta, joita tutkin kahdella eritavalla saadakseni monipuolisen näkemyksen tutkimukseni tueksi. Lisäksi erityyppisten tapausten valinnalla halusin varmistaa tutkimukseni yleistettävyyden.

Yin [1994, s. 6] toteaa, että tapaustutkimus (engl. case study) vastaa tyypillisesti selittäviin kysymyksiin ”kuinka” ja ”miksi”. Yin [1994, s. 13] määrittelee tapaustutkimuksen empiiriseksi tutkimukseksi, joka tutkii todellisen elämän ilmiöitä tilanteessa, jossa ilmiön ja asiayhteyden rajat eivät ole täysin selvät. Hän jatkaa määritelmää, että tapaustutkimus selviytyy omaperäisistä tilanteista, joissa on kiinnostavia muuttujia enemmän kuin tietopisteitä, mutta kuitenkin vain yksi lopputulos. Tapaustutkimus tukeutuu yleensä useisiin todisteaineistoihin, jollaista tapaa kutsutaan triangulaatioksi [Yin, 1994, s. 13]. Cunningham [1997] tarkentaa, että käytettävän tutkimusmenetelmän sisällä voidaan käyttää useita tekniikoita tietojen keräämiseen ja tulkitsemiseen. Hän huomauttaa, että triangulaatio voi myös olla usean menetelmän avulla tehtävää tutkimusta, esimerkiksi kyselyiden ja tarkkailun yhdistämisellä. Yin [1994, s. 13] mainitsee myös, että tapaustutkimus hyödyntää aikaisempia tutkimuksia ohjaamaan tiedon keräämistä ja analysoimista.

Tässä tutkimuksessa olen edennyt Yin [1994, s. 20-27] esittämien vaiheiden kautta. Hän jaottelee tapaustutkimuksen tekemisen viiteen tärkeään osatekijään. Tutkimuskysymysten asettelu muodostaa tapaustutkimuksen perustan. Tutkimuskysymysten pohjalta mahdollisesti esitetyt väitteet ohjaavat tutkintaa oikeaan suuntaan. Tutkimuskohteen määrittely esittelee tapauksen, kertoo millainen tutkittava kohde on. Tietojen keräämisen jälkeen tutkimuksen on myös esitettävä päättelytapa, kuinka kerätyt tiedot liittyvät väitteisiin. Lopuksi tutkimuksen on esitettävä tutkimuksen tulosten arviointi- ja tulkintaperusteet.

5.1. Tapausten valinta ja esittely

Tapausten valinta on keskeinen tapaustutkimuksen vaihe. Tapaustutkimuksen suunnittelutilanteessa vaihtoehtoina ovat yhden tapauksen tai monen tapauksen asetelmat [Yin, 1994, s. 38-39]. Yin jatkaa, että tapausta valittaessa on tehtävä myös päätös yhden tai useamman analyysiyksikön käyttämisestä tutkimuksen tekemisessä. Yhtä analyysiyksikköä käyttävää tutkimusasetelmaa nimitetään *holistiseksi* (engl. holistic) ja useampaa *sulautetuksi* (engl. embedded). Useamman tapauksen valintaa voidaan perustella esimerkiksi halulla varmentaa edellisestä tapauksesta saadut tulokset [Järvinen ja Järvinen, 2000, s. 83]. Mikäli tutkittava kohde koostuu useista analysoitavista yksiköistä, kuten organisaatiotutkimuksessa palavereista ja rooleista, voidaan hyödyntää esimerkiksi survey-tyyppistä lähestymistapaa [Yin, 1994, s. 41-44,51].

Valitsin tutkimusasetelmaksi monen tapauksen orientaation, koska haluan varmentaa teoreettisen viitekehysten soveltuvuuden ja tarkentaa ensimmäisen tapauksen tuloksia ja johtopäätöksiä. Koska tutkimusongelmani ja organisaatio-orientoituneen lähestymistavan käyttäminen tietovaraston suunnitteluun koostuu useista eri osatekijöistä, on tutkimuksessani myös useita analyysiyksiköitä.

Valitsin tutkimuksen kohteeksi kaksi tietovarastointiprojektia yrityksestä, jossa työskentelen. Valinnan perusteena oli projektien laajuus, pitkäkestoisuus ja monimuotoisuus sekä ajankohtaisuus ja tuoreus. Halusin myös erityyppiset tapaukset, koska näin teorian yleistämiselle on paremmat perusteet. Rajoitin tapausten kohteiden valinnan vain työskentelemääni yritykseen, koska tutkimusongelmani lähtökohta tuli omista ja työtoverieni kokemuksista. Tapausten valintaa tukivat myös saatavissa oleva monipuolinen materiaali ja tutkimusaineisto. Lisäksi käytettäessä kyselytutkimusta ovat läheiset työtoverit edullinen lähtökohta tutkimusaineiston helpompaan saatavuuteen, koska tällöin on suurempi todennäköisyys saada vastauksia, kuin jos kyselyn tekisi ventovieraille. Esimerkiksi Järvinen ja Järvinen [2000, s. 162] muistuttavat, että tunnetusti kyselytutkimuksen luotettavuuden heikkoutena on heikko vastausprosentti.

Olen kuvannut valitsemieni tapausten tutkimukseni kannalta keskeiset piirteet taulukossa 3. Näistä piirteistä käy ilmi, että tapauksessa yksi, yrityksessä A, tietovarastointi on keskeinen osa organisaation toimintaa. Tämä ilmenee tietovaraston kypsyydessä, käytön laajuudessa ja toimintatavoissa. Ensimmäinen tapaus on luonteeltaan prototyyppi tai pilottitapaus, jonka tarkoituksena on kuvailla organisaatio-orientoituneen lähestymistavan toteutumista tietovarastointiprojektin toteutuksessa. Pilottitapaus voi monen tapauksen tutkimuksessa auttaa täsmentämään tutkimussuunnitelmaa [Yin, 1994, s. 52]. Ensimmäinen tapaukseni ei kuitenkaan ole selkeästi pilottitapaus, jollaista ei yleensä raportoida, vaan se on tasavertainen tapaus toisen tapauksen rinnalla. Kuvailevan tapaustutkimuksen käyttämisestä puoltaa esimerkiksi projektin tekijöiden vähäinen määrä, jolloin survey-tyyppisen tutkimuksen

arvo saattaa olla heikko. Lisäksi projektin työntekijät ovat olleet projektissa kauan mukana, jolloin omalle työlle on voinut tulla sokeaksi, jolloin työtavat ja näkemykset ovat urautuneet.

Toinen tapaus, yritys B on luonteeltaan monessa suhteessa ensimmäisen tapauksen vastakohta (taulukko 3). Tietovarastointi on yritys B:lle uusi toimintamalli ja tämä ilmenee tietovaraston melko vähäisenä käyttämisenä, varsinkin organisaation työntekijöiden lukumäärään nähden. Mahdollisesti myös tietovaraston arvoa ja hyötyä ei ole täysin ymmärretty ja tämä näkyy esimerkiksi vanhojen tapojen puutteellisenä poisoppimisena. Tapaus kaksi on tutkimuslähestymistavaltaan selittävä ja pyrin tapauksen avulla vahvistamaan tutkimukseni teoriaa.

Kategoria	Alikategoria	Tapaus 1, yritys A	Tapaus 2, yritys B
Yritys	Liiketoiminnan luonne	valmistava teollisuus, logistiikka, jakelu	kuljetusala
	Tulosityksiköiden lkm.	noin 10	5
	Työntekijöiden lkm.	3000	12500
	Kansainvälisyys	kansainvälinen	kansainvälinen
Tietovarasto-järjestelmä	Loppukäyttäjien lkm.	400 - 500	100
	Kypsyys	9 vuotta	1 vuosi
	Tietolähteiden lkm.	8	10
	Sovellusalueiden lkm.	10	4
	Tietovaraston koko	85 GB	10 GB
	Paikallisvarasto vs. tietovarasto	Tietovarasto	Tietovarasto

Kategoria	Alikategoria	Tapaus 1, yritys A	Tapaus 2, yritys B
	Kehitys- metodologian pääpiirteet	Alussa kohteittain IT- vetoisesti, myöh. iteratiivinen, pieniä muutoksia ja yksi suurempi versiopäivitys. Tällä hetkellä asiakas hoitaa pääasiallisen kartoituksen sekä priorisoinnin ja toimittaja hoitaa teknisen määrittelytyön.	Aloitettiin raporttien määrittelyllä ja tämän jälkeen melko suoraviivainen koko tietovaraston toteutus. Tarpeet kartoitettu pääasiassa teknisen toteutuksen jälkeen ja vanhaa raportointi- järjestelmää mukailten.
	Raporttien lukumäärä	1000 raporttia, 50 tietokuutiota	200 raporttia, 3 tietokuutiota
Yrityksen tietovarasto- organisaatio	Työntekijöiden lkm.	20, joista 4 teknisiä	10 aktiivista, noin 35 osallistujaa kaikkiaan
	Organisaation piirteet	Liiketoimintavetoi- nen, dynaaminen, kasvava ja muuttuva organisaatio	Jäykkä, muutosvastarintainen, vanhoillinen
Toteuttava organisaatio	Työntekijöiden lkm. suurimmillaan	5	9

Taulukko 3. Tapaustudkimuksen kohteina olevien tietovarastointihankkeiden profiilit.

5.2. Materiaali ja tiedonkeruumenetelmät

Käyttämäni tiedonkeruumenetelmät vaihtelevat tutkittavasta tapauksesta riippuen. Hyödynsin tutkimuksessani sekä metodologista että tutkimusaineistoon liittyvää triangulaatiota. Käyttämäni tiedonkeruumenetelmät olivat kysely, osallistuva havainnointi ja projektien kirjalliseen materiaaliin tutustuminen sekä jossain määrin tekstien keruu. Useisiin tiedonkeruumenetelmiin päädyin, koska halusin monipuolista tietoa. Edellä mainitut tiedonkeruumenetelmät valitsin, koska niiden avulla pystyin parhaiten hyödyntämään projekteista saatavan informaation sekä projektien osallistujien sekä oman osaamiseni.

Dokumentaatio on yksi monista tapaustudkimuksen todistelähteistä. Ja se käsittää esimerkiksi muistioita, pöytäkirjoja ja hallinnollisia dokumentteja.. Käytettävissäni oleva dokumentaatio käsitti kummankin tapauksen osalta projektien teknisen ja liiketoiminnallisen dokumentaation sekä lisäksi suunnittelu- ja määrittelydokumentteja sekä muistiinpanoja. Arvioitaessa dokumentaation lähdeluotettavuutta vahvuuksiksi

nousevat esimerkiksi pysyvyys, tarkkuus ja yksityiskohtaisuus sekä laaja kattavuus. Heikkouksina voidaan nähdä esimerkiksi heikko saatavuus, kokonaisuuden epätäydellisyys ja tekijöiden mahdolliset näkymättömät painotukset ja mieltymykset. [Yin, 1994, s. 80-81]

Kyselyn valitsin koska, kyselyä suositetaan tiedonkeruumenetelmänä silloin, kun tiedusteltavia asioita on vähän, mutta vastaajia suhteellisesti paljon. Etuna on vastaajien vastausajankohdan vapaavalintaisuus. Kyselyn heikkoutena nähdään mahdollinen kysymysten ymmärtämisen vaikeus ja menetelmän kertaluontoisuus. [Järvinen ja Järvinen, 2000, s. 155]

Havainnoinnilla pääsin hyödyntämään omaa päivittäistä työtäni. Havainnoinnissa tietoja kerätään siten, että tutkija kirjaa huomioita ylös [Järvinen ja Järvinen, 2000, s. 162-163]. Järvinen ja Järvinen mainitsevat, että havainnointi eroaa haastattelusta siten, että tutkija näkee tutkimuskohteen omasta viitekehystänsä. He katsovat, että osallistuva havainnointi laajentaa ja konkretisoi tutkittavaa ilmiötä. Yin [1994, s. 80] muistuttaa, että tutkijan ennakoasenteet ja mieltymykset vaikuttavat tutkimuskohteen arvioimiseen.

Ensimmäisen tapauksen materiaalina käytin projektin aikana tuotettua dokumentaatiota (kuten teknistä kuvausta) ja keväällä 2006 projektin erään ison vaiheen päätteeksi suoritettua arviointi- ja palautetilaisuuden ylöskirjattuja huomioita. Dokumentaatio on kerätty kevään 2007 aikana. Lisämateriaalina hyödynnän projektin aikana käytyjä sähköpostikeskusteluja asiakkaan ja projektin toteuttavien henkilöiden kanssa. Lisäksi osallistuin ensimmäisessä tapauksessa kohteena olleen projektin toteutukseen, joten osallistuvalla havainnoinnilla ja arvioinnilla on merkittävä rooli tapauksen tiedonkeruumenetelmänä. Projektissa mukanaoloaikani kesti noin kevästä 2004 syksyyn 2006 ja vajaan vuoden ajan projektin loppupuolella kirjasin satunnaisesti ylös huomioita tätä tutkimustani varten. On kuitenkin huomioitava, että tämä projekti oli työurani ensimmäinen tietovarastointiprojekti.

Toisen tapauksen materiaali keskittyi kyselyllä suoritettuun tiedonkeruuseen. Käyttämäni kyselylomake oli osoitettu tässä tietovarastoprojektissa mukana olleille toimittajaorganisaation työntekijöille. Lisäksi käytössäni oli projektissa tuotettua dokumentaatiota, joka sisälsi jokaisesta olemassa olevasta dokumenttityypistä esimerkkikappaleen. Dokumentaation rooli tutkimuksessani oli tarjota vertailukohtaa kyselytutkimukseen. Dokumentaation keruuhetki oli keväällä 2007, kun projekti oli aikaisemmin vuoden alussa saatu päätökseen. Lisäksi projektin päätteeksi pidettiin vapaamuotoinen palaute- ja kehitystilaisuus, jonka keskustelut ja huomiot kirjattiin ylös. Nämä huomiot muodostavat tapauksen toisen tärkeän tutkimusaineiston.

Suoritin kyselytutkimuksen kevään 2007 aikana. Kyselylomake koostui avoimista kysymyksistä ja kysymykset oli strukturoitu eri osioihin. Jokainen kysymysosio koostui kartoittavasta tai kuvailevasta kysymyksestä sekä arvioivasta kysymyksestä.

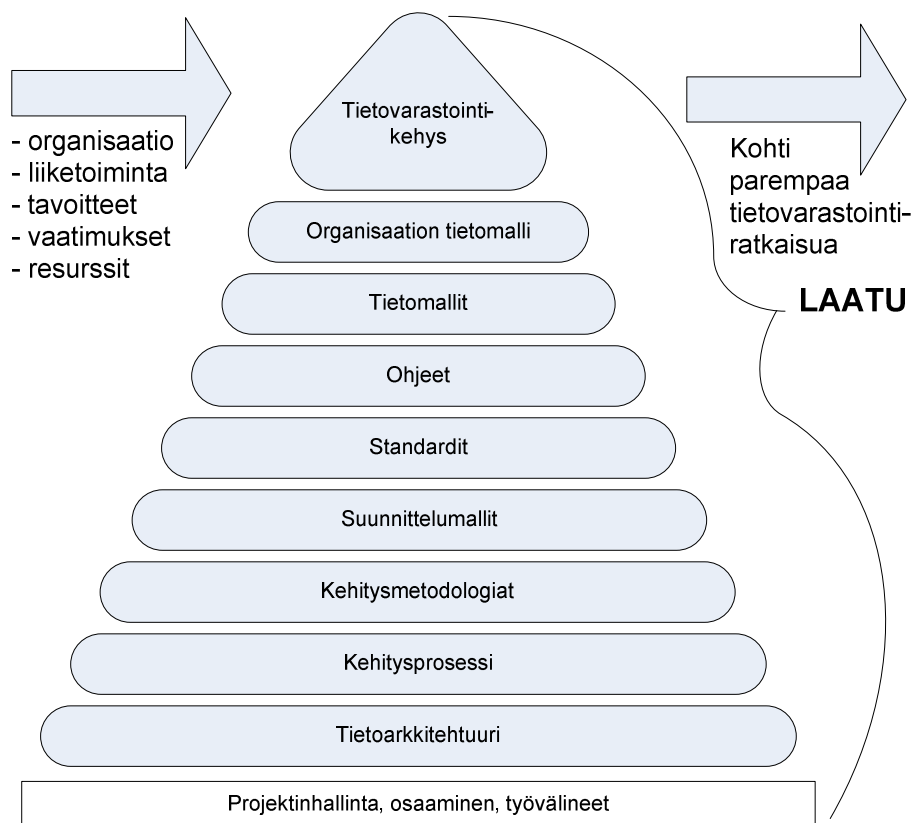
Kyselylomakkeen mukana meni sähköpostitse saatekirje, jossa kerrottiin esimerkiksi tutkimuksen tarkoitus. Kysely lähetettiin kaikille projektissa mukana olleille toimittajaorganisaatiomme työntekijöille. Kaikkiaan kyselylomakkeita lähetettiin 12 kappaletta ja tutkimukseen vastasi 10 työntekijää. Vastanneista kaksi olivat projektipäälliköitä ja loput kahdeksan olivat raportointi- ja tietovarastoasiantuntijoita, joilla oli vastuullaan erilaisia tehtäviä ja liiketoiminta-alueita tietovaraston suunnittelusta ja toteutuksesta.

5.3. Tutkimusasetelma ja analyysiyksiköt

Organisaatio-orientoituneen lähestymistavan soveltamisen tavoitteena tietovaraston kehitysprojektissa on kokonaisvaltaisen tietovaraston kehittämistapa ja tietovarastointiratkaisun hallittavuuden parantaminen tietovarastointikehystä käyttämällä. Lähestymistavan peruspiirteet käyttämiselle olen kerännyt kuvaan 25. Näistä saadaan tapaustutkimuksen analysointiyksiköt, jotka ovat

- tietovarastointiprojektin laatutavoitteiden asettaminen
- tietoarkkitehtuurin vaikutus projektin lopputulokseen
- tietovaraston kehitysprosessin käyttämisen vaikutus
- tietovaraston suunnittelussa käytetyt metatiedot, standardit ja ohjeet
- tietovaraston määrittely ja kehittäminen eri määrittelynäkökulmat huomioiden
- tietovaraston suunnittelumallien hyödyntäminen
- tietovaraston tietomallien käyttäminen mallintamisen apuvälineinä.

Tapaustutkimusten tavoitteena on selvittää jokaisen edellä mainitun tietovaraston suunnittelun ja kehitysprosessin osa-alueen vaikutusta tietovaraston suunnittelu- ja toteutusvaiheisiin tutkielmani organisaatio-orientoituneessa viitekehyksessä.



Kuva 25. Kokonaisvaltaisen organisaatio-orientoituneen lähestymistavan perusosat tietovaraston suunnittelussa.

Kuvassa 25 olen esittänyt tietovaraston suunnittelun organisaatio-orientoituneen lähestymistavan lähtötilanteen, tietovarastointihankkeen kiinteän perustan, lähestymistavan keskeiset osat ja tahtotilan eli tavoitteen. Esittelen seuraavaksi lähestymistavan jokaisen keskeisen piirteen teoreettisesta viitekehyksestä vedettyjen johtopäätösten perusteella.

Aluksi on selvitettävä tietovaraston rooli organisaatiossa. On määriteltävä tavoite, jokin liiketoiminnallinen ongelman ratkaisu, johon tietovarastoratkaisulla pyritään. Jotta osataan muodostaa sovelias työskentelytapa tietovaraston suunnitteluprosessiksi, on lisäksi asetettava selkeät tavoitteet tietovarastototeutukselle. Tavoitteiden ja yhteisen päämäärän asettamisella käyttäjäorganisaatio, sekä toimittajaorganisaatio pystyvät täysipainoiseen työskentelyyn.

Tietoarkkitehtuurin hyödyllisyys on todennettu sekä tietovaraston loogisen tietomallin että organisaation tietomallin luonnissa. Tietoarkkitehtuuri määrittää yritystasoisien tietostrategian, joka pohjustaa tietovaraston tietojen keräämistä operatiivisista järjestelmistä. Tietoarkkitehtuuri määrittää tietojen sijainnin organisaation tietojärjestelmäkentässä ja tulevaisuuden käyttömahdollisuuksia. Tietoarkkitehtuurin eräänä tavoitteena voidaan pitää keskitetyn tiedon (engl. master data) hallintaa.

Tietovaraston kehitysprosessin on oltava vaiheittainen ja liiketoimintalähtöinen. Kehitysprosessin keskeisintä osaa eli käsitteellistä mallia työstetään organisaation laajuisesti yhteistyössä liiketoimintaihmissen kanssa. Suunnitteluvaiheessa hyödynnetään sekä osittavaa että kokoavaa lähestymistapaa.

Metatietojen hallinnassa on huomioitava kaikki järjestelmän vaiheet. Dokumentoinnin ja ohjeistuksen on katettava sekä käyttäjien että toteuttajien tarpeet. Erityistä painotusta on käytettävä liiketoimintatietojen dokumentointiin. On suositeltavaa käyttää työkalua metatietojen hallintaan.

Tietovaraston määrittely käynnistyy liiketoimintavaatimusten keräämisestä. Myös organisaation operatiivisten järjestelmien tietojen analysointia suositellaan yhdistettäväksi vaatimusten keruuseen. Tietojenkeruun toteutustapana suositetaan liiketoimintaihmissen haastatteluita. Liiketoiminnan prosesseista ja toiminnasta koostetaan malleja, jotka hyväksytään tietovaraston kohdeorganisaation liiketoimintaihmisillä.

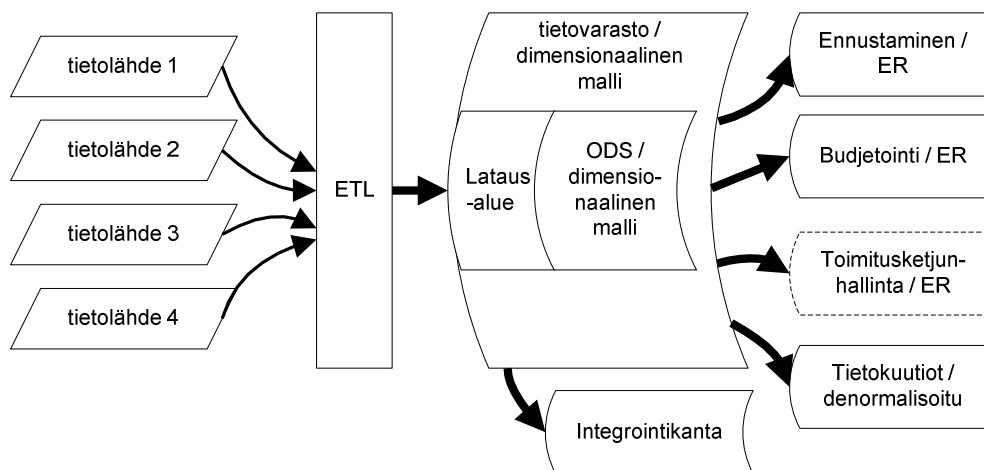
Tietovaraston suunnittelussa kannattaa hyödyntää ja yhdistää sekä ER-mallintamisen että dimensionaalisen mallintamisen hyviä puolia. Käsitteiden mallintamiseen ja luokitteluun kannattaa panostaa. Lopputuloksena on tarpeesta riippuen graafisesti kuvattu organisaation ja/tai tietovaraston tietomalli.

Tietomallin on kuvattava ja luokiteltava tietovaraston käsitteet moniulotteisesti. Tietovaraston mallintamisen ehdotetaan tapahtuvan käsitteellisen mallintamisen avulla yhdessä liiketoimintaihmissen kanssa. Mallin laajuuden määrittämisessä, hyväksyttämässä sekä informoinnissa on huomioitava koko organisaatio.

5.4. Tapaus 1: yritys A

Tapauksen tutkimisen tarkoituksena oli selvittää, kuinka yritys A:n nykyisessä tietovarastointiratkaisussa ja tietovaraston kehitystavoissa on hyödynnetty organisaatio-orientoitunutta kehitystapaa ja yrityksen tietomallia. Vertailukohtina käytän soveltaen kappaleessa 5.3 esiteltyjä yrityksen tietomallin piirteitä.

Yritys A toimii valmistavan teollisuuden alalla ja on yrityskooltaan mitattuna suuri. Yrityksen liiketoiminta on kansainvälistä ja omaan tuotekehitykseen perustuvaa. Yrityksessä nojaututaan vahvasti tietotekniikkaan ja yrityksen liiketoiminnan sydän on toiminnanohjausjärjestelmä (engl. Enterprise Resource Planning, ERP), jolla yritys ohjaa esimerkiksi tuotteidensa valmistusprosessia. Lisäksi muita tietojärjestelmiä on useita. Yrityksellä on verkkopalvelu kaupankäyntiä varten ja lisäksi yritys on verkostoitunut, integroitunut tietotekniikan avulla toisiin saman alan toimijoihin sekä asiakkaisiin. Yrityksen hallinnassa olevaa tietovarastoa hyödynnetään moniin tarkoituksiin, esimerkiksi teollisen valmistuksen seurantaan, päätöksentekoon, ennustamiseen. Yrityksen tietovarastointiratkaisu on kuvattuna kuvassa 26.



Kuva 26. Yritys A:n tietovarastointiratkaisu.

Yritys A:n tietovarastoratkaisun alkuperäisen version on tehnyt toinen toimittaja ja nykyisin käytössä olevat ETL- sekä raportointityövälineet ovat vaihdettu uudempiin teknologioihin tai päivitetty uudempiin versioihin. Tästä syystä historian painolastia toteutuksessa on erittäin paljon. Vanha dokumentointi oli Word- ja Excel-pohjaista ja hyvin hankalasti käytettävää yritys A:n tarpeisiin. Osittain Yritys A:n liiketoiminnan kiivasluontoisuudesta johtuen projektissa ei ole saatu riittäviä resursseja dokumentaation sekä tehtävienhallinnan nykyaikaistamiseksi. Lisäksi Yritys A:n tietohallintoa ei ole onnistuttu vakuuttamaan tarvittavien panostusten välttämättömyydestä, joten kehityksen hallinnan kannalta tilanne pahenee jatkuvasti. Erilaiset suoritettavat pienet korjaustoimenpiteet tulkittiin vain tilapäisiksi ja ongelmien siirtämiseksi tulevaisuuteen. Suurin ongelma tietovaraston käytössä oli sen käyttäminen sekä operatiiviseen eli taktiseen, että strategiseen raportointiin, kuitenkin pääpainon ollessa taktisessa raportoinnissa. Tämä ilmeni esimerkiksi monien uusien ominaisuuksien toteuttamisessa, joissa haluttiin vain raportoida tarkemmin toiminnanohjausjärjestelmän tietoja. Eli tietovarastoa käytettiin operatiivisen järjestelmän raportoinnin korvikkeena, koska suoraan operatiivisen järjestelmän päälle tehtävä raportointi olisi ollut liian kallista ja monimutkaista.

Keväällä 2006 suoritettiin Yritys A:n tietovarastointiprojektin toimittajaosapuolen projektihenkilöstön kesken palautetilaisuus, jossa kartoitettiin projektin erään ison vaiheen onnistumista ja parannustoimenpiteitä. Projektin työntekijöiden vastaukset olivat tehtävästä riippumattomasti hyvin yhdenmukaiset. Projektia kuvattiin sekavaksi, hitaaksi ja erittäin laajaksi. Arvioitaessa kommunikointia sisäinen kommunikaatio todettiin erinomaiseksi, mutta asiakkaan ja erityisesti asiakkaan liiketoiminta- sekä teknisten osaajien kanssa käyty kommunikaatio, erityisesti palaute, todettiin puutteelliseksi tai olemattomaksi. Tämä saattoi johtua monesta syystä, mutta tärkeimpinä pidän liiketoimintaosaajien päivittäisen työn aiheuttamaa kiirettä ja kiinnostuksen puutetta tietovaraston ja raportoinnin kehittämissä projekteissa. Vaikka

tietovarasto ja sen päälle rakennettu raportointi onkin liiketoiminnalliselta kannalta hyödyllisiä, niin liiketoiminta ei välttämättä pidä niitä kriittisinä.

5.4.1. Tietovarastointiprojektin laatutavoitteet

Yritys A:n tietovarastointihankkeelta on koko sen ajan, kun toimittajana on ollut Yritys X, puuttunut selkeät ja konkreettiset tavoitteet ja päämäärät, joihin tietovarastointi- ja raportointiratkaisulla on pyritty. Projektin aikana konkreettisin liiketoimintatavoite näytti mielestäni olleen tuotannon toiminnan seuraaminen laatua ja toimitusmääriä raportoimalla. Realistisia laatutavoitteita tietovarastolle ja sen sisältämille tiedoille ei projektissa määritelty lainkaan. Tästä on mielestäni aiheutunut esimerkiksi ongelmia virhetilanteiden selvittelyyn, koska Yritys A:n organisaatio ei ole määrittänyt toimintamalleja virhetilanteiden varalle eikä jakanut organisaation sisäisiä, järjestelmä- ja liiketoimintavastuita sekä järjestelmän toteuttajaosapuolen kanssa sovittavia vastuita.

5.4.2. Tietoarkkitehtuuri

Yrityksellä A ei ole käytössään yrityksenlaajuista tietoarkkitehtuuria. Järjestelmätasoinen yrityksen integraatioarkkitehtuurikuvaus on kuitenkin toteutettu neljä vuotta sitten. Käytännössä asiakkaan tietohallinto-organisaatiosta puuttuu tahto konsernitasoisen näkemyksen toteuttamiseksi, eli laadittu integraatioarkkitehtuurisuunnitelma on jäänyt osittain keskeneräiseksi. Esimerkiksi tietojen vastuujärjestelmien määrittäminen ja ylläpitoprosessien dokumentointi on suunnittelemta. Tulkitsin tilanteen kuitenkin niin, että liiketoiminnallista painetta tietoarkkitehtuurin laatimiseksi ei ole olemassa. Toisaalta, tietoarkkitehtuurin laatiminen voisi auttaa yrityksen liiketoiminnallisten ongelmakohtien ja mahdollisten tietoteknisten tehostetoimenpiteiden hahmottamisessa.

5.4.3. Tietovaraston kehitysprosessi

Tietovaraston kehitystyö hoidettiin täysin yrityksen ulkopuolisin voimin. Toteutuksen päättäntävalta oli käytännössä organisaation ulkopuolisella integraatio- ja tietovarastotoimittajalla. Tästä seurasi, että näkemys tulevaisuudesta puuttui ja tietojen omistajia ei ollut määritelty, joten määritysten mukana tuleva vastuu hämärtyi.

Uusien ominaisuuksien lisääminen tai muutostarpeiden toteuttaminen tietovarastoon on ollut liiketoimintalähtöistä, mutta pääosin teknisten yksityiskohtien toteuttamista. Esimerkiksi tehtävänä on ollut lisätä uusi toiminnanohjausjärjestelmän yksittäinen tieto poimittavaksi tietovarastoon. Pyynnöt ovat tulleet liiketoimintaihmisiltä, mutta yleensä ei kerrottu, miksi työ piti tehdä, eli mikä oli liiketoiminnallinen peruste työn toteuttamiseksi. Esimerkkinä tällaisesta yksityiskohdasta oli erään yksittäisen tiedon, toiminnanohjausjärjestelmän tietokannan taulun sarakkeen, lisäys tietovarastoon alkuperäisiä tietoja analysoimatta ja merkitystä selvittämättä. Toteutus tehtiin ja muutaman kuukauden kuluttua huomattiin, että osa tiedoista hylättiin virheellisinä

tietoina. Toteuttaja oli päätelty tietosisällön olemassa olevien tietojen perusteella ja siksi ominaisuuden koko oli rajoitettu, koska sen havaittiin sisältävän pelkästään pieniä numeroita. Virheellinen tieto sisälsi kuitenkin pitkän merkkijonon ja asia ratkaistiin siten, että asiakkaan tietohallinnon edustaja pyysi kasvattamaan attribuutin kokoa. Ei selvitetty, mikä oli tiedon tarkoitus tai miksi se oli tietovarastoon poimittu, eikä käyty läpi lähtöjärjestelmän attribuutin sisältöä ja sen merkitystä. Asiakkaan liiketoimintaihmisia ei myöskään koulutettu tietovaraston ja raportoinnin käyttämiseen, joten tehtäväpyyntönä saattoi olla dimensioattribuutin lisäys faktatauluun, jotta liiketoimintaihmissen raportti olisi helpompi toteuttaa.

Kehitysprosessia tai suunnitteluprosessia ei ollut kuvattu lainkaan, joten jokainen uusi toteutustehtävä saatettiin suorittaa eri tavoin. Joitakin tehtävien vaiheita saatettiin hyväksyttää tai tarkistuttaa Yritys A:n liiketoimintaosaajilla tai tietohallinnon työntekijöillä, mutta tästä ei ollut mitään sovittuja käytäntöjä. Toisaalta käytettyjen työkalujen puutteellisuus ja heikkous metatietojen ja projektin elinkaaren hallintaan ja erityisesti usean version ja ympäristön hallintaan aiheutti jatkuvasti ongelmia samanaikaisessa kehityksessä sekä projektin ylläpitovaiheessa.

5.4.4. Metatiedot, standardit ja ohjeet

Projektissa ei käytetty juurikaan ohjeistusta eikä standardeja. Yhteisiä pelisääntöjä ei sovittu ja kirjattu ylös tai niiden järjestyttä ei mietitty, josta seurasi määrittelyä, toteuttamiseen sekä dokumentointiin kirjavia toimintatapoja. Osaksi tämä johtui projektiryhmän haluttomuudesta, koska työtehtäviä oli ilman lisätyötäkin riittävästi ja osittain siitä syystä, että toiminnalle ei välttämättä olisi saatu Yritys A:n rahoitusta. Ohjeistusten ja standardien laadintaankin voisi mielestäni soveltaa iteratiivista menetelmää, jolloin tehtävä ei tuntuisi niin raskaalta.

Ongelmallisina asioina kevään 2006 palautetilaisuudessa todettiin esimerkiksi, että liiketoimintatietämys kohdealueesta sekä lähtöjärjestelmän käyttötapojen tunteminen helpottaisi työskentelyä. Lisäksi sekä lähtöjärjestelmien, että tietovaraston dokumentoinnin todettiin olevan puutteellista. Tämä koski sekä teknistä että liiketoiminnallista dokumentaatiota. Osaksi tästä syystä esimerkiksi tehtävien siirto henkilöltä toiselle vaatii syvällistä kohdeaiheen asiantuntemusta. Tämä dokumentoinnin hajanaisuus, puutteellisuus sekä ajantasaisuuden puuttuminen käy ilmi, kun tutkii ratkaisusta olemassa olevaa dokumentaatiota. Dokumentaatiota on tehty monin tavoin hyödyntäen tietämyksenhallintaa, Word- ja Excel-muotoisia dokumentteja, tehtävienhallintajärjestelmää, tietokannan suunnittelutyökalua sekä suoraan tietokantaan dokumentoiden. Keskitetty, johdonmukainen dokumentointi sekä teknisen ja liiketoiminnallisen näkemyksen yhdistäminen samaan tietämyksenhallintajärjestelmään olisi ensiarvoisen tärkeää. Projektin dokumentaation ajantasaisuuden hallitseminen olisi myös ratkaistava. Erityisesti asiakasorganisaation tietohallinnon työntekijät ovat toivoneet parempaa dokumentaatiota. Toisaalta asiakasorganisaatio ei ole ymmärtänyt

dokumentoinnin nykytilaa ja sen haltuun ottamisen tarvitsemaa panostusta. Myös asiakkaan ja toimittajan sähköpostikeskustelun perusteella ilmeni selkeä tarve tilanteen ottamiseksi haltuun, koska monet pyynnöt olivat selvityspyyntöjä tai esimerkiksi tietojen etsintää dokumentoinnista asioiden oikeellisuuden selvittämiseksi. Lopputuloksena asiakas yleensä ihmetteli, mistä metatiedot voisi löytää, ja toimittajaosapuoli joutui selvittämään asian suoraan toteutuksesta.

5.4.5. Tietovaraston määrittely ja kehittäminen

Tietovaraston määrittelyssä ei lainkaan otettu huomioon liiketoimintaprosesseja analysointinäkökulmasta, toisin sanoen määrittäminen tehtiin hyvin teknisesti. Esimerkiksi dokumentointia tutkimalla ilmenee seuraus liiketoimintaprosessien ja sääntöjen puuttumisesta, jolloin dokumentaatioissa on kuvattu suoraan monimutkainen sql-kyselykielen lause tietyn toiminnan tekemiseksi, mutta ei kertomalla liiketoimintasääntöä luonnollisen kielen avulla. Määrittelyvaiheen pääasiallisena tavoitteena olikin yleensä selvittää, kauanko työhön menee aikaa eli, paljonko määriteltävän ominaisuuden toteuttaminen tulisi maksamaan. Koska liiketoimintaosaajat olivat hyvin vähän tekemisissä toteuttajien kanssa määrittämis- ja kehitysvaiheissa, toteuttajat joutuivat itse ratkaisemaan esimerkiksi tietojen esitysmuodot. Mikäli asiaan halusi tarkennusta, yritys A:n tietohallinnon vastaus oli, että toteutetaan niin kuin ennenkin, vaikka tuloksena olisi lisää sekavuutta toteutukseen.

5.4.6. Tietovaraston suunnittelumallit

Tietojen mallintaminen suoritettiin suoraan loogisen relaatiomallintamisen avulla käyttäen Oraclen Designer-työkalua ja mallintamisen lopputuloksena oli relaatiotietokantakaavioita. Tällöin käsitteellinen mallinnusvaihe ohitettiin kokonaan. Tietovaraston tietomallin suunnittelussa ei hyödynnetty liiketoimintaihmissä juuri lainkaan ja vain korkean tason vaatimukset tulivat liiketoimintaihmisiltä. Suunnittelussa keskityttiin suoraan yksityiskohtaisuuksiin, eikä tulevaisuutta, liiketoiminnan muuttumista arvioitu riittävästi. Liiketoiminnan muuttuessa on tällöin epäselvää, mitä toiminnallisuutta tai tietorakenteita tarvitsee muuttaa tietovarastoon. Ongelmallisena koin myös liiketoimintasääntöjen hajauttamisen eri järjestelmäkerroksiin, kuten ETL-prosessin eri vaiheisiin, suoraan tietokantatoteutuksiin, raportointiin, tietokuutioihin ja muihin liiketoimintatiedonhallintajärjestelmiin. Osittain tämä johtui käytetyistä työvälineistä, jotka eivät tukeneet yhtenäistä mallinnusta. Ongelman voisi ratkaista siten, että liiketoimintasäännöt mallinnettaisiin pelkästään ETL-työvälineen lataus- ja muunnosprosessien avulla. Tällöin tietovarastoon mallinnetut tiedot muodostaisivat valmiin pohjan erilaisille analysointi- ja raportointitarpeille.

5.4.7. Tietovaraston tietomallit

Organisaation tietomallin puuttumisen seurauksena oli se, että yrityksen kattava näkemys liiketoimintatiedoista puuttui. Tällä hetkellä Yritys A:ssa käytetään pääasiassa vain yhdestä lähtöjärjestelmästä peräisin olevaa tietoa. Tämä järjestelmä on toiminnanohjausjärjestelmä. Tietoa on mallinnettu sisäisten yksikköjen pohjalta ja mallintamisen toteutuksen seurauksena on muodostanut eräänlaisia paikallisvarastoja tietovaraston sisälle. Tästä mallinnustavasta on seurauksena ollut päällekkäisiä asioita. Konsernitasoinen mallintaminen saattaisi korjata nämä suunnitteluvirheet. Konsernitasoisen tietomallin toteuttaminen paljastaisi myös mahdolliset nykyisiin järjestelmiin liittyvän vajaakäytön.

Käsitteet eivät olleet yksiselitteisiä ja niitä ei kuvattu ja dokumentoitu lainkaan. Tietovaraston suunnittelijat eivät myöskään hyväksyttäneet käyttämiään termejä ja käsitteitä liiketoimintaihmisillä. Toisaalta myöskään liiketoimintaihmiset eivät näistä olleet kiinnostuneita. Nimeäminen oli epäyhtenäistä ja epäselvää. Nimeämisessä ei mietitty standardoitujen nimien käyttöä yrityksen laajuisesti, vaan on käytetty suoraan pääasiassa yhden lähtöjärjestelmän, toiminnanohjausjärjestelmän terminologiaa. Myös tietovaraston tietojen yhteismitallistaminen puuttuu osittain, koska käytössä on esimerkiksi eri mitta- ja rahayksiköitä tietovaraston sisällä.

Dimensioiden hyödyntäminen tietovarastossa oli minimaalista. Esimerkiksi käsitteiden attribuutteja toistettiin tietomallissa, jolloin sama attribuutti saattoi olla monella käsitteellä. Esimerkkinä tällaisesta toteutuksesta käy organisaatiodimensio, eli yhtiö-, yksikkö-, varasto- ja paikkahierarkiassa ei viitata pelkästään alimpaan dimension tasoon vaan myös dimension ylempiin tasoihin. Tehokkuusvaatimuksia ei voinut pitää syynä, koska tehokkuusvaatimuksia ei ollut määritelty ja käsitteiden sisältämät tietomäärät eivät olleet niin suuria, jotta ne olisivat aiheuttaneet tehokkuusongelmia. Dimensioita ei myöskään mallinnettu kattamaan koko organisaation keskeisiä käsitteitä, joten vain peruskäsitteet, kuten asiakas ja tuote on mallinnettu dimensioina.

5.4.8. Muita huomioita

Yrityksen järjestelmäintegraatiossa operatiivista tietovarastoa ei ole hyödynnetty juuri lainkaan, vaan sen lisäksi on olemassa rinnakkainen järjestelmä, joka on tehty käytännössä yhtä käyttötarkoitusta, tilauksien käsittelyä ja integrointia varten. Samalla tätä integrointijärjestelmää kuitenkin hyödynnetään myös muussa tarkoituksessa. Operatiivista tietovarastoa ei myöskään käytetty hyväksi yksityiskohtaisessa tai tosiaikaisessa raportoinnissa. Se toimi vain osittain normalisoituna tiedon keräysväylänä tietovarastoon. Alun perin operatiivista tietovarastoa oli tarkoitus käyttää esimerkiksi organisaation keskeisten käsitteiden yhdistämiseen, kuten usean lähtöjärjestelmän asiakkaiden yhdistämiseen ja jakamiseen eteenpäin muihin järjestelmiin, jolloin

operatiivinen tietovarasto toimisi organisaation alkuperäisten tietojen yhdistämiskanavana.

Positiivisena puolena havaitsin, että liiketoimintatiedon johtaminen on käynnistynyt budjetoinnin ja ennustamisen osalta. Nämä tietovarastoon läheisesti sekä liiketoimintatiedonhallintaan liittyvät järjestelmät ja tietoalueet ovat saaneet korkean tason liiketoimintatukijan ja suunnitelman tietojen yrityksen laajuiseen hyödyntämiseen.

5.4.9. Organisaatio-orientoitunut lähestymistapa ja organisaation tietomallit tapauksessa 1

Projektissa käytetty kehitysprosessi sekä suunnittelumenetelmät olivat hyvin teknisiä eikä liiketoimintatietoja mallinnettu riittävästi. Organisaation tietomallia ei käytetty, joten tietoja toistettiin, eikä yrityksellä ollut selkeätä näkemystä hallitsemistaan tiedoista.

5.5. Lähtökohdat tapaustutkimukselle 2

Kuvailevassa tapaustutkimuksessani kävin läpi erään yrityksen tietovarastointihankkeen. Teoreettisen viitekehyksen perusteella laadittua organisaatio-orientoinutta lähestymistapaa ja teorian pohjalta laadittujen tapaustutkimuksen analyysiyksiköiden käsittelyä peilaan seuraavaksi eräässä toisessa tapauksessa, Yritys B:n tietovarastointihankkeessa.

Toteutin toisen tapaustutkimuksen kyselynä, jolloin sain mahdollisimman kattavan osallistumisen tutkimukseeni. Kysely suunnattiin kaikille Yritys B:n tietovarastointiprojektin toteuttajaorganisaation jäsenille, riippumatta heidän rooleistaan ja tehtävistään projektissa. Tällöin sain myös luotettavamman otoksen projektin toiminnasta, koska usea projektiin osallistuja vastasi samoihin kysymyksiin. Kyselyn vastauksia peilaan projektissa tuotettuun dokumentaatioon sekä täydennän vastauksia keväällä 2007 pidetyn projektin palautetilaisuudessa kerrottujen ajatusten perusteella.

Kyselytutkimukseni kysymykset ovat kuvattuna liitteessä 1. Kysymykset on jaoteltu seitsemään osioon taustakysymyksiin, projektin laatutavoitteiden käsittelyyn, tietoarkkitehtuurin käsittelyyn, kehitysprosessin kuvaamiseen, tietovaraston määrittelyyn ja kehittämiseen, tieto- ja suunnittelumallien käsittelyyn sekä metatietoihin, standardeihin ja ohjeisiin. Lisäksi vastaajat saivat antaa vapaamuotoista palautetta tietovarastointihankkeisiin liittyen.

5.6. Tapaus 2: yritys B

Yritys B:n tietovarastointiprojekti oli monin tavoin vastakkainen kuin yritys A:n tietovarastointiprojekti, kuten taulukosta 3 selviää. Toisaalta oli mielenkiintoista verrata kehitysprosessien ja organisaatiokulttuurien erilaisuuden vaikutusta tietovaraston kehittämiseen. Yritys B:n tietovarastointiprojektissa oli myös näkyvissä historian

painolastia, koska yhtenä määrittysten perusteena käytettiin vanhaa raportointijärjestelmää ja esimäärittelyä oli aikaisemmin ollut tekemässä toinen yritys, joten puhtaasti tyhjältä pöydältä ei tässäkään projektissa päästy aloittamaan.

5.6.1. Tietovaraston tavoitteiden asettaminen

Yritys B:n tietovarastoratkaisun ylimalkaiset laatutavoitteet olivat asetettu hankkeen tarjouksessa ja sopimuksessa. Sen jälkeen määrittelyä ei ollut tarkennettu, eli varsinaista laatutavoitteiden määrittelyä ei suoritettu, mikä johtui osaksi projektin nopeasta käynnistymisestä. Monelle oli kuitenkin epäselvää, että projektille olisi lainkaan asetettu minkäänlaisia laatutavoitteita tai ainakaan niistä ei projektin sisällä kommunikoitu. Tärkein toimittavan organisaation asettama laatutavoite oli toimittaa sopimuksessa sovittu toiminnallisuus riippumatta siitä, vastasiko tämä asiakkaan tarpeita. Monet kommentoivat myös epärealistisen aikataulun aiheuttamia ongelmia ja vastuiden puuttumista, jolloin esimerkiksi asiakasorganisaatio saattoi toimittaa erilaista aineistoa suunnittelu, toteutus- ja testivaiheissa.

Mittareita laatutavoitteille ei ollut käytettävissä, eikä niitä haluttu projektin aikana kehittää. Tämä olisi vienyt aikaa, joten ilmeisesti rahapoliittisista syistä kehitykseen ei haluttu uhrata aikaa. Parannusehdotuksina esitettiin esimerkiksi mitattavia tavoitteita, kuten suoritusnopeuden mittaamista. Eräs ehdotus oli laatutavoitteiden asettaminen sekä siirtotiedostojen toimittajille eli niin sanotuille kolmansille osapuolille sekä asiakkaan tuottamille määrittelyille.

Tärkeimmäksi projektin aloitusvaiheen tehtäväksi ehdotettiin yritys B:n nykytilan kartoitusta sekä tulevaisuuden tarpeiden analysointia. Tämä tarve ilmeni esimerkiksi siten, että asiakkaan tarpeita ei kyseenalaistettu, joten määrittelyssä keskityttiin nykyisen raportointiympäristön korvaamiseen. Keväällä 2007, projektin palautetilaisuudessa esitettiin, että asiakkaan koulutukseen on kiinnitettävä huomiota jo heti projektin aluksi. Tietovaraston käyttötapojen ja mahdollisuuksien opettaminen auttaisi projektin yhteisten tavoitteiden muodostamisessa.

5.6.2. Tietoarkkitehtuuri

Projektin asiantuntijoiden vastausten perusteella yritys B:llä ei ollut käytössään tietoarkkitehtuuria eikä sellaista myöskään suunniteltu projektin aikana. Vastauksissa todettiin, että tietovaraston koordinointi oli erittäin puutteellista tai se puuttui kokonaan, jolloin esimerkiksi eri liiketoiminta-alueiden tietoja ei saatu onnistuneesti yhdistettyä ja organisaatiolle ei pystytty miettimään tietoarkkitehtuuria nykyisten eikä tulevaisuuden tarpeiden näkökulmasta edes tietovaraston osalta.

Projektiin piti vastausten perusteella kuulua yrityksen yhteisen tietomallin suunnittelu, mutta yhtenäisestä tietomallista muodostui vain raportoinnin dimensiotietojen tallennuspaikka. Eräässä vastauksessa jatketaan, että tietojen välitystä tietomallista edelleen muiden järjestelmien käyttöön ei mietitty lainkaan. Osittain tästä

syystä asiakkaan ja toimittajan käsitykset tietoarkkitehtuurista tietojen latauksen ja integroinnin osalta erosivat toisistaan. Toiseksi syyksi esitettiin myös asiakkaan epärealistista kuvaa organisaation tietoarkkitehtuurista.

Suunnitteluvaiheessa eri lähtöjärjestelmistä pyrittiin löytämään yhteneviä tietoja ja suorittamaan näiden tietojen yhdistäminen dimensioiksi. Hyötyä tästä ei kuitenkaan erään tekijän mukaan saatu, koska dimensiot oli rajattu liiketoiminta-alueen mukaan ja millään raportilla ei käytetty eikä ilmeisesti tulla käyttämään koko organisaatiota koskevia dimensioita. Tähän esitettiin ratkaisuksi tietovaraston käyttötarkoitusten selkeyttämistä ja kuvaamista.

Eräs vastaaja totesi, että yksittäisistä lähtöjärjestelmistä olisi pitänyt kiinnittää eniten huomiota merkittävimpään järjestelmään. Tähänkin tietoarkkitehtuuri olisi mielestäni tuonut ratkaisun, koska siitä olisi käynyt selväksi organisaation kriittiset ja monimutkaisimmat järjestelmät. Eräessä toisessa vastauksessa korostettiin lähtöjärjestelmien tietojen laadun heikkouden ja näiden järjestelmien toimittajien heikon osaamisen merkittävää vaikutusta projektiin.

Projektissa olisi erään vastauksen perusteella voinut olla yksi tai useampia asiantuntijoita, jotka olisivat osanneet koordinoida koko tietoarkkitehtuurin mukaisen kehityksen ja tiedonjalostusprosessin lähtöjärjestelmistä raportointiin. Lisäksi olisi huomioitava yhtenäisen tietoarkkitehtuurin vaatima koulutustarve projektin tekijöille. Nyt havaitsin vastauksista, että kaikki eivät ymmärtäneet, mitä sillä tavoitellaan ja mikä on sen hyötyarvo. Tämä saattoi näkyä myös tietovarastointihankkeen korkeimmalla tasolla, koska hanketasolla työryhmien päätöksiä hylättiin jatkuvasti. Projektin lopputuloksena syntyneitä tietovaraston arkkitehtuuria pidettiin kuitenkin melko onnistuneena.

5.6.3. Tietovaraston kehitysprosessi

Selkeätä kehitysprosessia ei tekijöiden mielestä ollut lainkaan käytössä vaan se muotoutui projektin kuluessa. Eräs kuvaus oli seuraava:

”Tehtiin mitä ehdittiin, siinä järjestyksessä, kun tuntui parhaalta.”

Monet vastaajat mainitsivat käytössä olleen asiakkaan vaatimuksesta vesiputousmallin, jossa kaikki toiminnallisuus toimitettiin kerralla. Tällainen kehitysprosessi ei vastaajien mielestä soveltunut lainkaan liiketoimintatiedonhallinnan projekteihin. Yksi vastaaja totesi kuvaavasti kehitysprosessin lopputuloksesta näin:

”Vesiputousmallilla voidaan varmistaa, että asiakas saa kerralla kokonaisen järjestelmän, joka ei vastaa hänen tarpeitaan.”

Vesiputousmallin huonoina puolina esitettiin esimerkiksi aikataulullisesti myöhäistä järjestelmien integrointia sekä raportointivaatimusten ja tietovaraston tietomallin vaikeaa yhdistämistä. Yhdistämisen vaikeus johtui lähtöjärjestelmien sisällön

tutkimisesta vasta raporttimääritysten jälkeen, jolloin esimerkiksi puuttuvista tiedoista tuli ongelmia, koska niihin ei määrittelyssä ollut otettu kantaa.

Soveliaampana mallina ehdotettiin vaiheittaista toiminnallisuuden käyttöönottoa. Prosessin tarpeellisuudesta projektihenkilöt olivat samaa mieltä. Pidettiin tarpeellisena, että käytössä olisi kehitysprosessi, jossa olisi määritelty kaikki tarvittavat työvaiheet kronologisessa järjestyksessä. Yksi vastaaja kuitenkin mainitsi, että kun projektin myöhemmässä vaiheessa kehitystyö ja käytetyt työkalut olivat asiakkaalle tutumpia, vesiputousmallikin toimi paremmin. Projektin palautetilaisuudessa esitettiin, että tietovaraston tietomallin eri osa-alueiden toteuttaminen olisi tehtävä peräkkäin, jolloin jo saman projektin sisällä voidaan hyödyntää opittua.

Eräs vastaaja kuvasi käyttämänsä kehitysprosessin melko tarkasti, vaikka mainitsikin, että kehitys muotoutui tällaiseksi vasta työskentelyn aikana. Siinä lähdettiin asiakkaan tarvemäärittelystä, joita vastasivat raporttimääritykset, ja jatkettiin tietotarveanalyysin kautta tietokannan suunnitteluun. Tämän jälkeen asetettiin lähtöjärjestelmille toimitusvaatimukset. Seuraavaksi lähtöjärjestelmien tiedot yritettiin integroida ja normalisoida, jonka jälkeen tiedot siirrettiin normalisoimattomaan tähtimalliin raportointia varten.

Tärkeitä kehitysvaiheita jäi kiireen takia väliin ja esimerkiksi ETL-prosessien suunnittelu ja toteutus yhdistyivät yhdeksi työvaiheeksi. Esitettiin myös näkemys, että suunnitteluratkaisut olivat lyöty lukkoon jo sopimusvaiheessa, joten projektin tekijöiden toimintaa oli rajoitettu jo ennen projektin alkamista, jolloin asiakkaiden oikeiden tarpeiden täyttäminen hankaloitui. Kehitysprosessiin ehdotettiin myös määritysten tarkentamista, jotta alkuperäisiin esimäärittelydokumentteihin ei olisi tarvinnut turvautua enää projektin myöhemmissä vaiheissa sekä päätöksentekopisteitä prosessin eri vaiheisiin. Esimerkiksi testausvaiheiden perään olisi pitänyt saada asiakkaan hyväksyntä, mutta toisaalta eräässä vastauksessa mainittiin dokumenttien katselmoinnin olleen liian haasteellista asiakasorganisaatiolle.

Monen eri toimittajien sekä heidän hallitsemien ympäristöjen kanssa toimiminen olisi vaatinut hyvin määritellyt prosessit sekä aikatauluista ja laadusta kiinnipitämisen. Projektiin olisi myös vaadittu tehokkaampaa kommunikointia sisäisten vastuualueiden välillä erityisesti suunnittelun ja toteutuksen osalta, koska vastauksista käy ilmi, että vastuunjako ei ollut selkeä.

Projektissa toteutettua prototyypiratkaisua kritisoitiin vain teknologiaorientoituneesta lähestymistavasta. Prototyypin toteutuksen aikana olisi yhdessä asiakkaan kanssa voitu harjoitella työtapoja ja tutustuttaa asiakas uusiin työkaluihin. Prototyyppien käyttämistä pidettiin myös hyvänä tapana tutustua lähtöjärjestelmään.

5.6.4. Tietovaraston määrittely ja kehittämisen

On mielenkiintoista havaita, että tietovaraston määrittelyä kuvattiin tavoiteohjatuksi, koska asiakasorganisaation talousosasto oli määritellyt pääasialliset tavoitteet. Toisaalta muut liiketoiminta-alueet ajoittain vastustivat näitä tavoitteita. Useat vastaajat kuvasivat kuitenkin käytettyä määrittelytapaa tieto-ohjatuksi, koska oli olemassa oleva järjestelmä, joka piti korvata. Tapaa kuvattiin onnistuneeksi, koska asiakkaalla oli riittävä tietämys lähtöjärjestelmistä ja liiketoiminnan vaatimukset eivät olleet monimutkaisia. Haastavana pidettiin kuitenkin lähtöjärjestelmän ja tietovaraston vaatimusten yhteensovittamista ja tietojen muokkausprosessin jakamista lähtöjärjestelmien ja tietovaraston kesken.

Kehittämiseen ehdotettiin iteratiivista lähestymistapaa, jossa liiketoiminta-alue kerrallaan olisi suoritettu määrittely ja toteutus. Tätä perusteltiin esimerkiksi sillä, että jokainen yritys B:n liiketoimintaorganisaatio olisi ollut paremmin valmistautunut ja asennoitunut omaan tehtäväänsä. Määrittelemällä tietovaraston arkkitehtuurille käyttötapaukset, olisi tietovaraston perustan toteutukseen saatu erään vastaajan mukaan realistinen arkkitehtuuri kehitystyön pohjaksi.

Useat vastaajat korostivat, että projektissa keskityttiin liiaksi toteuttamaan vanha raportointijärjestelmä uudella teknologialla, jolloin määrittelyn pohjana olivat vanhat raportit. Eräs vastaaja kertoi määrittelytyön toimineen tietovaraston osalta hyvin, mutta raportoinnin osalta huomattavasti heikommin. Parempana lähestymistapana pidettiin liiketoimintatiedonhallinnan kehittämistä yrityksessä ja vasta tämän jälkeen suunniteltaisiin järkevät raportoinnin käytötavat. Projektin loppupuolella määrittelyä tehtiin hybridimallia mukailen, hyödyntämällä sekä tavoite- että tieto-ohjattua lähestymistapaa, jolloin oli selvitetty asiakkaan liiketoimintatarpeet ja tutkittu miten olemassa olevilla tiedoilla päästiin mahdollisimman lähelle tarpeita. Projektia kuvattiin hyvin määrittelyintensiivisiksi, mikä johtui projektin rahoitusmallista. Toisaalta määrittelyvaiheisiin panostaminen antaisi erään vastaajan mukaan paremman mahdollisuuden esimerkiksi henkilövaihdoksiin.

Projektin määrittely- ja suunnitteluvaiheissa tehtiin päätöksiä, jotka erään vastaajan mielestä vaikuttivat koko toteutuksen onnistumiseen. Ilmeisesti projektissa määriteltiin ja hyväksyttiin raporttien määrittelyvaiheessa myös siirtotiedostot, jolloin ETL-suunnittelun yhteydessä havaittiin, että niihin tarvitaan suuriakin muutoksia. Tässäkin tapauksessa tietojen siirtäminen määrittelytyöryhmästä toteuttajalle oli ollut heikkoa. Tästä syystä voisikin mielestäni olla parempi käyttää useaa eri lähestymistapaa tietovaraston määrittelyssä, esimerkiksi sekä jäsentävän, että osittavan käyttämistä suunnittelun työkaluina. Eräänä ratkaisuna kokonaisvaltaisen määrittelyn toteuttamiseksi ehdotettiin kahden määrittelijän sekä tietovarastoasiantuntijan, että raporttoijan käyttämistä määrittely ja suunnitteluvaiheissa. Vastauksissa ehdotettiin myös useamman projektipäällikön käyttämistä määrittely- ja toteutusvaiheissa. Yksi

vastaaja ehdotti myös lähtöjärjestelmän asiantuntijan käyttämistä tietovaraston määrittelyvaiheessa.

Lisäksi todettiin, että tietovarastointikehyksen olemassaolo olisi voinut helpottaa työntekoa, koska erään vastaajan sanoin – kaootinen malli – oli henkisesti raskas. Esimerkiksi formaalin muutoksenhallinnan käyttäminen olisi selkiyttänyt määrittelytyötä, jolloin tilanteet olisi voitu jäädä testausta varten. Nyt ongelmina nähtiin esimerkiksi testattavan aineiston muuttuminen kesken testauksen sekä muutoksien kumuloituminen projektin työmääriin. Eräs vastaaja ehdotti, että lähtöjärjestelmistä tuotavien tietojen formaateista olisi myös pitänyt pitää kiinni, koska periksi antamalla liikaisten tietojen määrä lisääntyi jatkuvasti ja projektin osapuolien vastualueet sekoittuivat. Eräs asiantuntija esitti myös yksityiskohtaisen toimintasuunnitelman käyttöönottoa, jolloin kehitysvaiheet olisivat olleet etukäteen selvillä. Muutamassa vastauksessa suositeltiin lisäksi dokumenttipohjien ja tarkistuslistojen käyttämistä kaikissa projektin vaiheissa.

5.6.5. Tietovaraston tieto- ja suunnittelumallit

Tietovaraston suunnittelussa hyödynnettiin käsittekartoitusta, käsitteellistä mallintamista ER-mallin avulla sekä dimensionaalista mallintamista, joista kuitenkin vain dimensionaalinen mallintaminen esiintyi melkein jokaisessa vastauksessa. ER-mallin avulla generoitiin tietokantamalli. Käsitteet mallinnettiin korkealla tasolla sekä kaikki dimensiot ja faktat luettelointiin. Tietomallin keskeinen sisältö kerättiin määrittelyvaiheen työkokousten pohjalta. Tietomallia ei kytketty lähtöjärjestelmien tietoihin, mutta vastaajien mukaan tämä ei aiheuttanut ongelmia. Toisaalta esitettiin kommentti, että yhtenäinen mallintaminen suunnitellun prosessin mukaan olisi helpottanut tiedonjakamista projektin jäsenten välillä. Eräessä vastauksessa samoin kuin palautetilaisuudessa mainittiin, että yhteistä tietovaraston suunnitteluvaihetta ei ollut lainkaan. Erityisesti tietovarastointi- ja raportointitekijöiden välillä koettiin olleen tiedonjakokatkoksia varsinkin sen jälkeen, kun osa tekijöistä oli siirtynyt jo uusiin tehtäviin. Olettavasti toteutuksen kuvaaminen dokumentaatioon oli kuitenkin jäänyt riittämättömäksi. Tietovaraston tietomalleissa olisi pitänyt erään vastaajan mielestä varautua paremmin suuriin tietomääriin esimerkiksi lisäämällä valmiiksi summatasoista tietoa.

5.6.6. Metatiedot, standardit ja ohjeet

Metatietoja kertyi käytettyihin työvälineisiin eli käytettyyn ETL-työkaluun Oracle Warehouse Builderiin sekä raportointityövälineisiin. Toisaalta eräs vastaaja kertoi, että metatietojen hallintaan ei ollut kunnollisia työkaluja. Projektiin liittyvää liiketoimintasanastoa kerättiin käytettyyn tietämyksenhallintajärjestelmään, mutta koska nämä asiat eivät olleet välttämättömiä projektin toteutuksen kannalta, niiden ylläpito jäi kesken ja niitä ei päivitetty toteutustyön ja testauksen aikana. Eräs vastaaja kuitenkin

huomautti, että liiketoimintasanaston kuvaus olisi pitänyt antaa edes osittain asiakkaan tehtäväksi.

Tietämyksenhallintajärjestelmän ja tehtävienhallintajärjestelmän käyttöä kehitettiin edellyttäen, että niillä on selkeä rakenne ja yhdessä sovitut käyttötavat. Dokumentointi hoidettiin kuitenkin pääosin Word- ja Excel-dokumentteihin liiketoiminta-aluekohtaisesti. Metatiedoista kerättiin tietojen tasotietoa, mutta ei tietojen tyyppejä tai laatua. Osittain määrittelytyötä tehtiin tietämyksenhallintajärjestelmään, johon kuvattiin esimerkiksi tietomallien käsitteet. Muutamassa vastauksessa kuitenkin mainittiin, että käytössä ei ollut yleistä liiketoimintasanakirjaa. Termistön kerääminen suositeltiin toteutettavaksi määrittelyvaiheessa heti projektin aluksi, jolloin muutoinkin on avattava käsitteitä. Liiketoimintakäsitteistön päivittäminen ja käytön ohjeistamista ehdotettiin palautetilaisuudessa. Tämä käsitteistö olisi pitänyt vielä hyväksyttää asiakkaalla, todetaan eräissä vastauksessa.

Word-pohjaista dokumentaatiota parempana ratkaisuna eräs vastaaja piti keskitettyä dokumentointiratkaisua, jossa kaikki käsitteet ja niiden väliset suhteet olisivat olleet kuvattuina. Word-dokumentteja pidettiin yleisesti vaikeasti ylläpidettävänä. Eräänä parannusehdotuksena esitettiin metatietojen keruuta järjestelmän tietosisällön tilasta, esimerkiksi taulujen ETL-latausten tilanteesta, yhdistettynä www-pohjaiseen käyttöliittymään, jolloin käytössä olisi helposti lähestyttävät tietovarastojärjestelmän tilatiedot. Vastauksissa esitettiin myös kritiikkiä projektidokumentaation toteutuksen ajankohdasta, koska suurin osa projektidokumentaatiosta tehtiin vasta projektin toteutustyön päätyttyä. Dokumentaation päivitys oli esimerkiksi saattanut unohtua tai sen paikkansapitävyys saattoi olla kyseenalaista. Tarvittava toimitettava dokumentaatio ei myöskään ollut selvillä ennen projektin toteutusvaiheen päättymistä. Muutama vastaaja arvioi, että olisi pitänyt kerätä tietoa laskentasäännöistä ja allokointiperusteista, koska niiden puute tai heikko kuvaus aiheutti toteutusvaiheessa ylimääräistä selvitystyötä. Erityisesti tästä olisi hyötyä silloin, kun toteuttaja on eri henkilö kuin määrittelijä. Nämä määrittelyt olisi myös pitänyt hyväksyttää asiakkaalla.

Eräs vastaaja oli sitä mieltä, että metatietojen kerääminen jäi kokonaan toteuttamatta. Todennäköisesti hän tarkoittaa, että kukin suunnittelija keräsi oman osa-alueen osalta joitakin tarvittavia tietoja, jolloin yhtenäisiä ja kattavaa metatietoa tietovaraston osalta ei ollut olemassa. Ideaalidokumentaatioksi esitettiin sekä kyselyn vastauksissa että palautetilaisuudessa dokumentaation linkitystä toteutukseen ja dokumentaation tuottamisen automatisoimista sekä pohjadokumenttien käyttämistä.

Tietokannan nimeämiskäytäntöihin vastaajat olivat pääosin tyytyväisiä, vaikka ennen projektia olisi pitänyt olla valmiina enemmän ohjeistusta. Kaksi osallistujaa oli lisäksi sitä mieltä, että nimistandardejakaan ei ollut käytössä. Tämä ilmeni esimerkiksi siten, että toisten toteuttamiin ETL-prosesseihin ei haluttu koskea. Yleisesti toivottiin helposti löydettävää ja keskitettyä dokumentaatiota sekä dokumenttimallien valmistelua

ja käytön koulutusta ennen niiden käyttöönottoa. Ohjeistuksen ja standardien osalta ehdotettiin, että kaikki käytännöt pitäisi sopia ETL-toteuttajien ja raportintekijöiden yhteisissä palaverissa. Tällä vältettäisiin tekijöiden omien käytäntöjen ja toimintatapojen syntyminen. Lisäksi palautetilaisuudessa kannatettiin metatietolähtöistä tietokantadokumentointia, jolloin käytössä olisi esimerkiksi jokin pelkästään metatietojen hallintaan tarkoitettu työväline.

Kävin läpi projektin dokumentaation ja totesin sen olleen vastaajien kuvaaman kaltainen. Melkein kaikki tietovaraston kehitysprosessin elinkaaren vaiheet olivat kuitenkin huomioitu dokumentaatiossa läpi edes lyhyesti. Toisaalta dokumentaatio oli hyvin ylimalkaista, kiireessä tehtyä ja sisälsi huomattavasti toistoa. Suunnitelmissa toimittavan organisaation ohjelmistoprojektikokemuksen vaikutus näkyi selkeästi ja lisäksi esimerkiksi ohjeistus keskittyi pääasiassa testaamisen toteuttamiseen. Dokumentaatiossa korostuivat erityisesti teknisyys ja työvälineorientoituneisuus. Sanastoa ja liiketoiminnan laskusääntöjä oli kuvattu melko kevyesti. Johtopäätöksenä totesin, että keskitetty dokumentoinnin rakenteen kuvaavan tietämyksenhallinnan käyttäminen ratkaisisi monia dokumentaation ongelmia, jolloin vaihetuotteet sekä ohjeistukset syntyisivät pakon sanelemana. Asiakkaan liiketoimintaan liittyvien metatietojen kuvaamiseen ja dokumentoimiseen pitäisi kiinnittää enemmän huomiota.

Palautetilaisuudessa esitettiin mielenkiintoinen kommentti organisaation oppimisesta. Siinä organisaation todettiin oppivan liian hitaasti. Metatietojen kerääminen, dokumentaation ja ohjeistuksen helppo, kattava sekä ajantasainen saatavuus johtaisi jo opittujen asioiden tehokkaampaan levittämiseen, jolloin myös oppiminen paranisi.

5.6.7. Vastaajien yleisiä kommentteja tietovarastoinnista

Eräs vastaaja piti tietovarastointiprojekteja tekijän kannalta epäkiitollisina, koska lähtöjärjestelmien tietojen virheet ilmenevät usein tietovarastossa, mutta virheitä ei kuitenkaan korjata lähtöjärjestelmiin vaan tietovarastointiratkaisuun pitää toteuttaa erikoiskäsittelyä, jolloin toteutuksesta tulee virhealttiimpaa. Ennen projektin aloitusta on siksi mielestäni sovittava selkeät pelisäännöt, toimintatavat ja vastuut sekä tietojen omistajuus, jotta virhetilanteissa on selkeät toimintatavat. Vastauksissa toivottiin myös työvälineiden kehittyvän ja erityistä huomiota esitettiin tarvittavien metatietojen hallintaan ja avoimuuteen.

Yksi vastaajista oli sitä mieltä, että tehokkaan projektinhallinnan edellytys on sopiva ja riittävä projektin rajaus sekä muutostenhallinnan käyttäminen. Myös projektin johdon etäisyyttä työryhmiin kritisoitiin liian suureksi, vaikka työryhmät onnistuivatkin toimimaan itsenäisesti. Eräs huomio oli, että ohjelmistoprojekteissa hyväksi havaitut menetelmät eivät suoraan toimi tietovarastointiprojekteissa. Tietovarastointiprojekti erosi erään vastaajan mukaan perinteisestä ohjelmistoprojektista monilla osa-alueilla, esimerkiksi määrittelyn, versioinnin ja testauksen osalta. Tähän lisättiin myös, että

projektipäällikön olisi hyvä tuntea käytettävät teknologiat riittävällä tarkkuudella. Myös toimittajavetoisuutta kritisoitiin ja asiakkaan raportointivastaajille olisi pitänyt mahdollistaa raporttien luonti omiin tarpeisiin olemassa olevista tiedoista.

5.6.8. Organisaatio-orientoitunut lähestymistapa ja organisaation tietomallit tapauksessa 2

Yritys B:n tietovaraston suunnitteluprosessia ja kehitysvaiheita vaivasivat samat ongelmat kuin yritys A:n tietovarastointiprojektia. Kehitys- ja suunnittelumenetelmät eivät olleet jäsenyneitä, eikä organisaation näkemyksiä huomioitu tai niitä ei onnistuttu määrittämään. Organisaation tietoa mallinnettiin kuitenkin melko hyvin kaikilla tiedontasoilla koko organisaatio huomioiden, vaikka tietoarkkitehtuuriin ei kiinnitettykään riittävästi huomiota.

Dokumentoinnin hallintaa ja tasoa pidin merkittävänä ongelmana vielä projektin päätyttyäkin. Monista eri seikoista johtuen tietovarastointiratkaisun kytkeminen osaksi liiketoimintatietoja jäi hyvin keskeneräiseksi.

Tapauksesta havaitsin useita projektin epäonnistumisen mahdollistavia syitä, mutta edes suuret lyhytaikaiset ongelmat eivät silti saisi peittää alleen projektin sellaisia laiminlyöntejä, joilla on pitkäaikaisia vaikutuksia. Suurimmat yksittäiset vaikutukset projektiin olivat asiakasorganisaation jäykät ja sopimusvetoiset toimintatavat. Projektissa korostui selkeästi, että on välttämätöntä, että kaikki hankkeessa mukanaolijat kulkevat kohti samaa päämäärää.

6. Tulokset ja havainnot

Esittelen tässä luvussa tapaustutkimusteni pohjalta tekemäni johtopäätökset sekä arvioin tutkielmani merkityksellisyyttä. Lopuksi pohdin tutkielmani havaintojen vaikutusta tietovarastointiprojekteihin sekä esitän ratkaisumalleja havaintojen tueksi.

Organisaation tietojen kokonaisvaltainen hyödyntäminen tietovaraston kehittämisessä ja mallintamisessa johtaa organisaation pitkän aikavälin kustannussäästöihin, koska tietovarastointi tukee strategisesti organisaation liiketoimintaa. Lisäksi organisaation liiketoiminnan tukeminen metatietoihin perustuvalla tietovarastolla mahdollistaa organisaation tehokkaamman oppimisen. Organisaatio-orientoitunut lähestymistapa tukee esimerkiksi organisaation tietojen uudelleen järjestelyä.

Organisaation tietomalli määrittää organisaation informaatiotarpeisiin perustuvan tietopääoman jakamisen yrityksen tietojärjestelmiin. Organisaation tietomalli toimii esimerkiksi tietovaraston dimensionaalisen mallintamisen pohjana ja pitkän tähtäimen liiketoimintatietojen mallina.

6.1. Päätelmät tapaustutkimuksien osalta

Tapaustutkimukseni tapauksissa oli samankaltaisia ongelmia esimerkiksi kehitysprosessin hallinnan osalta. Yritys B:n tapauksessa projektin sisällä oli vastauksissa havaittavissa ristiriitoja, joten projektissa käytetyt menetelmät ja toimintatavat olivat yksilöllisiä, eivät yhteneväisiä. Kummassakin tapauksessa näkyi tietovarastointiprojektien kokemattomuuden mukanaan tuomia ongelmia erityisesti tietovaraston ominaispiirteiden huomioimisessa sekä keskittyminen tekniikkaan liiketoiminnan sekä kehitys- ja suunnitteluprosessien kustannuksella.

Erityistä tarvetta näin usean suunnittelutavan käyttämisellä ja suunnittelun etenemisen seuraamisen tarkistuslistojen avulla, jotta suunnittelussa huomioidaan kaikki näkökulmat lähtötiedoista liiketoimintavaatimuksiin. Erityisesti kokonaisvaltainen näkemys puuttui pääosin kummankin projektin kehitysprojektin hallinnasta. Kummassakaan yrityksessä yrityksen tietoja ei myös hallittu kokonaisvaltaisesti ja pohdittu niiden strategista arvoa riittävästi. Sekä toimittaja- että asiakasosapuolia yhdistivät metatietojen ja dokumentoinnin arvon vähätteleminen, mikä näkyi heikkotasoisessa dokumentaatiossa ja ohjeistuksessa.

6.2. Tutkimuksen arviointi, luotettavuus ja jatkotutkimustarpeet

Aloitin tutkimukseni tekemisen tietomalleista ja organisaation tietojen ominaisuuksien sekä vaikutuksien tutkinnalla. Jatkoin aiheen käsittelyä organisaatio-orientoituneen lähestymistavan piirteiden tutkimisella, mutta hyvin nopeasti huomasin, että tutkimuskohteet eivät olleet irrallisia eivätkä itsenäisiä kokonaisuuksia, vaan siihen liittyi kokonaisvaltaisesti monia muita asioita ja toimintatapoja, joten käsittelemäni aiheen rajaaminen muodostui haasteelliseksi.

Teorian ja käytännön yhdistäminen oli haastavaa, koska yritysmaailma ja tieteellinen tutkimus ovat mielestäni verrattain kaukana toisistaan. Uskon, että tutkielmastani on kuitenkin hyötyä yrityksellemme ja voimme tutkielmani pohjalta jatkokehittää työskentelytapojamme, osaamistamme sekä ratkaisuja liiketoiminnan tarpeisiin.

Kyselytutkimuksen luotettavuus on mielestäni hyvä, vaikka projektin päättymisestä olikin useilla projektiin osallistuneilla melko paljon aikaa. Kulunut aika oli saattanut osittain haalistaa tapahtuneita ja koettuja asioita, mutta toisaalta ajan kulumisen saattoi myös vahvistaa joitakin yksityiskohtia ylitse niiden arvon. Muutama vastaaja myös kritisoi laatimieni kysymysten vaikeutta, moniselitteisyyttä sekä monien outojen termien käyttöä. En kuitenkaan kokenut itse tätä ongelmaksi, koska kysymykset olivat mielestäni muotoiltu siten, että varsinaiset kysymykset olivat yksinkertaisia, mutta ohjeet ja vastausvihjeet saattoivat olla vaikeita. Kysymyksiin saattoi siten vastata, vaikka ei olisikaan hyödyntänyt kyselylomakkeeseen kirjoitettuja avainsanoja.

Tutkielmani käsittelemä aihe antaa mahdollisuuden moniin jatkokysymyksiin, koska tietovaraston kehitys-, määrittely- ja mallinnusasiat riippuvat monista yksittäisistä tekijöistä. Lisäksi tietovarastoinnin heterogeeninen ja standardoimaton tilanne johtaa useisiin erilaisiin ongelmanratkaisutapoihin. Liiketoiminnan huomioimisen tarkempi tutkinta esimerkiksi yrityksen tietostrategiaa ja yritysmalleja tutkimalla voisi tuoda uusia näkökulmia tietovaraston rooliin yrityksen päätöksentekovälineenä. Konkreettisenä seuraavana vaiheena olisi laatia ohjeistus, dokumentaatio sekä käyttökelpoinen menetelmä esittelemäni organisaatio-orientoituneen lähestymistavan peruserätyksien soveltamiseksi yritysmailmaan. Erityisen mielenkiintoinen tämän tutkimuksen kannalta on Stefanovin [2006] tekeillä oleva väitöskirja, joka käsittelee tietovarastoinnin ja organisaation keskinäistä vuorovaikutusta. Hänen tavoitteenaan on tehostaa tietovarastoinnin ja organisaation suhdetta käsitteellisillä malleilla sekä tietojen tulkintaa liiketoiminnan metatietojen avulla.

6.3. Keskustelu

Tutkimuksen kohteena olleiden yritysten toiminnan perusteella totesin, että kohdeyritykset eivät olleet täysin oivaltaneet tietovaraston monipuolisia mahdollisuuksia. Kriittistä liiketoimintaa ohjaavat toiminnanohjausjärjestelmät vievät päättäjien pääasiallisen huomion ja tietovaraston merkitys voi hämärtyä. Yleensä tilanne on kuitenkin korjattavissa, jos yrityksessä tiedostetaan tietovaraston liiketoiminnallinen merkitys. Tietovaraston avulla voidaan jopa löytää uusia mahdollisuuksia liiketoiminnalle, mikäli yrityksen tietoarkkitehtuurille löydetään vastuuhenkilöt ja tukijat.

Tilanteen parantamiseksi ehdotan kehittämisprosessin sekä tietovarastointi- ja raportointiratkaisun tuotteistamista paketoituksi kokonaisuudeksi. Paketointi perustuisi prosessorientoituneeseen toimintamalliin, jossa olisi selkeästi määritellyt tavoitteet, toiminnot ja eri vaiheisiin liittyvät tuotteet. Tuotteistamiseen kuuluisi eri osa-alueita ja komponentteja kattaen projektin suunnittelu- ja toteutusvaiheet. Sisältönä olisi tietovarastointiratkaisun merkityksen kouluttaminen, asiakkaan liiketoimintaongelmien kartoittaminen ja esitys niiden ratkaisemiseksi tietovarastointiratkaisun avulla. Eräänä vaiheena voisi olla pilottiprojektin toteuttaminen ja esimerkiksi hyviksi havaittujen valmistyökalujen sekä aputyökalujen käyttäminen, joiden avulla ratkaisun toteuttaminen sujuisi hallitusti. Lisäksi paketointi sisältäisi valmiit ratkaisut tehtävien- sekä tietämyksenhallinnan tarpeisiin hyödyntämällä valmiita tarkoitukseen soveltuvia tuotteita. Esimerkiksi Confluence-tietämyksenhallintatuotteen avulla muodostettua projektin dokumentointirunkoa voisi käyttää jokaisen tietovarastointiprojektin metatietolähtöisen dokumentaation perustana. Perusteellinen kehitysprosessin ohjeistaminen ja dokumentoiminen tukisi kehitysprosessia ja asiantuntijoiden työntekoa. Asiakas voisi ostaa paketista asteittain eri osa-alueita. Tietämyksenhallintaratkaisulla toteutettaisiin keskitetty, reaaliaikainen kaikille

osapuolille avoin strukturoitu dokumentointi, joka yhdistäisi kehitysprosessin eri vaiheet sekä tietovaraston tiedot ja liiketoimintatiedot saumattomasti toisiinsa sekä toimisi projektin koko elinkaaren ajan käyttökelpoisena koulutusympäristönä ja osaamisen jakokanavana.

Kokonaisvaltaisen tietovarastointinäkömyksen edistämiseksi käytössä voisi esimerkiksi olla valmiit suunnitelmat erilaisiin tietoarkkitehtuurin avulla ratkaistavissa oleviin yleisiin liiketoimintaongelmiin. Näin asiakkaan tarpeet otettaisiin huomioon liiketoimintaongelmalähtöisesti, mutta näkökulma olisi kuitenkin organisaation kattava. Tällainen tietoarkkitehtuuriratkaisu huomioisi esimerkiksi operatiivisen ja taktisen raportointitarpeen operatiivisen tietovaraston avulla.

Kuvaamalla sovellettava kehitysprosessi tietovarastointiorganisaatiolle saadaan välitettyä selkeä kuva järjestelmän toteuttamisen vaiheista. Käytettävistä menetelmistä voidaan myös sopia yhteistyössä asiakkaan kanssa, mutta tällöin on oltava varovainen toimintatapojen vaikutuksista tietovarastointiprojektiin.

Hankkeissa olisi myös selvitettävä erillisen metatietojen kuvauskannan rakentamista todellisena vaihtoehtona metatietojen hallintaan. Myös keskitetyn metatietojen hallinnan työvälineen käyttöä kannattaisi harkita. Tämä on kuitenkin hankalaa, koska eri sovellukset ja sovellustoimittajat käyttävät erilaisia ratkaisuja ja standardeja metatietojen kuvaamiseen.

Käsittekartoitukseen olisi kiinnitettävä erityistä huomiota. Käsittekartoituksen ja mallintamisen apuvälineinä kannattaa käyttää esimerkiksi tutkielmassani esitettyjä tapoja luokitella ja taulukoida käsitteitä. Käsitteelliseen mallintamiseen liittyen pääasiassa loogisten tietojen kuvaamiseen käytettyä dimensionaalista kuvaamistapaa kannattaisi soveltaa myös käsitteellisen mallin esittämiseen, koska tällainen esitystapa auttaa liiketoimintaihmissä ja kehittäjiä paremmin hahmottamaan mallinnettava liiketoiminta-alue sekä raporttien pohjana oleva tietovaraston tietomalli.

Viiteluettelo

- [Adelman and Moss, 2000] Sid Adelman and Larissa T. Moss, *Data Warehouse Project Management*. Addison-Wesley Professional, 2000.
- [Andrighetto and Assogna, 2004] G. Andrighetto and Pierluigi Assogna, Why we understand each other: the role of the context. In: *Proc. 3rd CAiSE Workshops*, 2004, 293-296.
- [Arnott and Pervan, 2005] David Arnott and Graham Pervan, A critical analysis of Decision Support Systems research. In: *Journal of Information Technology*, 20, 2 (Jun. 2005), 67-87.
- [Ballard *et al.*, 1998] Chuck Ballard, Dirk Herreman, Don Schau, Rhonda Bell, Eunsang Kim and Ann Valencic, *Data Modeling Techniques for Data Warehousing*, IBM redbook. International Technical Support Organization, IBM Corporation, 1998. Available online at <http://www.redbooks.ibm.com/redbooks/pdfs/sg242238.pdf> (8.4.2007)
- [Böhnlein and Ende, 2000] Michael Böhnlein and Achim Ulbrich-vom Ende, Business process oriented development of data warehouse structures. In: Reinhard Jung and Robert Winter (eds.), *Data Warehousing 2000 - Methoden, Anwendungen, Strategien*. Physica-Verlag, Heidelberg, 2000, 3-21.
- [Elmasri and Navathe, 2000] Ramez Elmasri and Shamkant B. Navathe, *Fundamentals of Database Systems*, 3rd edition. Addison-Wesley, Reading MA, 2000.
- [Calvanese *et al.*, 1999] Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, Daniele Nardi and Riccardo Rosati, A principled approach to data integration and reconciliation in data warehousing. In: *Proc. of the Intl. Workshop on Design and Management of Data Warehouses*, CEUR Workshop Proceedings **16** (1999), Technical University of Aachen.
- [Calvanese *et al.*, 2006] Diego Calvanese, Luigi Dragone, Daniele Nardi, Riccardo Rosati and Stefano M. Trisolini, Enterprise modeling and data warehousing in Telecom Italia. *Inform. Syst.* **31**, 1 (Mar. 2006), 1-32. Elsevier Science Ltd., Oxford, UK.
- [Chaudhuri and Dayal, 1997] Surajit Chaudhuri and Umeshwar Dayal, An overview of data warehousing and OLAP technology. *SIGMOD Rec.* **26**, 1 (Mar. 1997), 65-74. ACM Press, New York, NY, USA.
- [Chen, 1976] Peter Pin-Shan Chen, The entity-relationship model—toward a unified view of data. In: *ACM Trans. Database Syst.* **1**, 1 (Mar. 1976), 9-36. ACM Press, New York, NY, USA.
- [CMS, 2007] U.S. Department of Health & Human Services, Centers for Medicare & Medicaid Services (CMS), Data administration. <http://www.cms.hhs.gov/DataAdmin/> (3.3.2007).

- [Counihan *et al.*, 2002] Ailish Counihan, Pat Finnegan and David Sammon, Towards a framework for evaluating investments in data warehousing. *Inform. Syst. J.* 12, 4 (2002), Blackwell Science Ltd, 321-338.
- [Cunningham, 1997] J.Barton Cunningham, Case study principles for different types of cases. *Quality and Quantity* 31, 4 (Nov. 1997), 401-423.
- [Darke *et al.*, 1998] Peta Darke, Graeme G. Shanks and Marianne Broadbent, Successfully completing case study research: combining rigour, relevance and pragmatism. In: *Inform. Syst. J.* 8, 4 (Oct. 1998), 273-290.
- [Eriksson ja Koistinen, 2005] Päivi Eriksson ja Katri Koistinen, Monenlainen tapaustutkimus. Julkaisuja 4. Kuluttajatutkimuskeskus, 2005. Saatavilla myös: http://www.ncrc.fi/files/4957/2005_04_verkkojulkaisu_tapaustutkimus.pdf (16.4.2007).
- [Galliers and Land, 1987] Robert D. Galliers and Frank F. Land, Viewpoint: choosing appropriate information systems research methodologies. In: *Commun. ACM* 30, 11 (Nov. 1987). ACM Press, New York, NY, USA, 900-902.
- [Gardner, 1998] Stephen R. Gardner, Building the Data Warehouse. *Comm. ACM* 41, 9 (Sep. 1998), 52-60.
- [Goodhue *et al.*, 1992] D. L. Goodhue, M. D. Wybo and L. J. Kirsch, The impact of data integration on the costs and benefits of information systems. *MIS Q.*, 16, 3 (Sep. 1992), 293-311.
- [Halpin, 2001] Terry Halpin, *Information Modeling and Relational Databases: from Conceptual Analysis to Logical Design*. Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, 2001.
- [Hannula and Pirttimäki, 2004] Mika Hannula and Virpi Pirttimäki, A cube of business information. *SCIP04 International Conference & Exhibit*, March 22-25, 2004, Boston, Massachusetts, USA.
- [Harding and Yu, 1999] J.A. Harding and B. Yu, Information-centred enterprise design supported by a factory data model and data warehousing. *Comput. Ind.*, 40, 1, (Sep. 1999), 23 - 36. Elsevier Science Publishers B. V.
- [Herrmann, 2004] Clemens Herrmann, Exploring the structural dimension of data warehouse organizations: results of a survey and implications. In: R. Meredith, G. Shanks, D. Arnott, S. Carlsson (eds.): *Proc. of the 2004 IFIP International Conference on Decision Support Systems (DSS2004)*. Monash University, 2004, 350-358.
- [Hoppenbrouwers *et al.*, 2005] Stijn Hoppenbrouwers, Henderik Alex Proper and Theo P. van der Weide, Understanding the requirements on modelling techniques. *Advanced Information Systems Engineering, 17th International Conference (CAiSE'05)*, Lect. Notes Comput. Sci., 3520. Springer, 2005, 262-276.

- [Hüsemann *et al.*, 2000] Bodo Hüsemann, Jens Lechtenböcker and Gottfried Vossen, Conceptual data warehouse design. In: *Proc. of the International Workshop on Design and Management of Data Warehouses (DMDW'00)*, CEUR Workshop Proceedings, **28** (2000), CEUR-WS.org, 6:1-6:11.
- [Inmon *et al.*, 1997] William H. Inmon, J. D. Welch and Katherine L. Glassey, *Managing the Data Warehouse*. Wiley, New York, NY, USA, 1997.
- [Inmon, 2002] William H. Inmon, *Building the Data Warehouse*, 3rd edition. John Wiley & Sons, 2002.
- [Inmon, 2005] William H. Inmon, *Building the Data Warehouse*, 4th edition. Wiley Publishing, Inc, 2005.
- [Jarke *et al.*, 1998] Matthias Jarke, Manfred A. Jeusfeld, Christoph Quix and Panos Vassiliadis, Architecture and quality in data warehouses. In: *Proc. of the 10th International Conference on Advanced Information Systems Engineering (CAiSE '98)*, Lect. Notes Comput. Sci., **1413** (1998). Springer-Verlag, 93-113.
- [Jarke, 2001] Matthias Jarke, Enterprise model-driven data warehousing. Elektronische Festschrift zum 60. Geburtstag von August-Wilhelm Scheer Source Wissensmanagement: das Magazin für Führungskräfte (Juli, 2001).
- [Jones and Song, 2005] Mary Elizabeth Jones and Il-Yeol Song, Dimensional modeling: identifying, classifying & applying patterns. In: *Proc. of the 8th ACM international Workshop on Data Warehousing and OLAP (DOLAP '05)*, (Nov. 2005). ACM Press, New York, NY, USA, 29-38.
- [Jukic, 2006] Nenad Jukic, Modeling strategies and alternatives for data warehousing projects. *Comm. ACM*, **49**, 4 (Apr. 2006), ACM Press, New York, NY, USA, 83-88.
- [Järvinen ja Järvinen, 2000] Pertti Järvinen ja Annikki Järvinen, *Tutkimustyön metodeista*. Opinpaja Oy, Tampere, 2000.
- [Kaldeich and Sá, 2004] Claus Kaldeich and Jorge Oliveira e Sá, Data warehouse methodology: a process driven approach. In: *Proc. of Advanced Information Systems Engineering, 16th International Conference (CAiSE 2004)*, Lect. Notes Comput. Sci., **3084** (2004). Springer, 536-549.
- [Kendle, 2005] Noreen Kendle, The enterprise data model. The Data Administration Newsletter (TDAN.com), Robert S. Seiner, Jul 2005. Available online at <http://www.tdan.com/i033hy04.htm> (7.3.2006).
- [Kimball and Ross, 2002] Ralph Kimball and Margy Ross, *The Data Warehouse Toolkit*, 2nd edition. Wiley, New York, 2002.
- [Lamb, 2001] Joanne Lamb, *Sharing best methods and know-how for improving generation and use of metadata*. In: Eurostat (ed.) Pre-Proc. *ETK/NTTS*, **1** (Jun 2001), 175-194. Available as http://webfarm.jrc.cec.eu.int/etk-ntts/Papers/final_papers/19.pdf. Checked 15 December 2006.

- [Long, 1998] Kathy Long, The enterprise data model: a key ingredient for successful data warehousing. The Data Administration Newsletter (TDAN.com), Robert S. Seiner, Jun 1998. Available online at <http://www.tdan.com/i005fe12.htm> (7.3.2006).
- [Loucopoulos, 2000] Pericles Loucopoulos, From information modelling to enterprise modelling. In: Sjaak Brinkkemper, Eva Lindencrona and Arne Solvberg (eds.), *Information Systems Engineering: State of the Art and Research Themes*. Springer, New York, NY, 2000, 67-78.
- [Martin, 1982] James Martin, *Strategic Data Planning Method*. Prentice Hall PTR, NJ, USA, 1982.
- [Martin *et al.*, 2004] Richard A. Martin, Edward L. Robertson and John A. Springer, Architectural principles for enterprise frameworks: guidance for interoperability. *International Conference on Enterprise Integration and Modelling Technology (ICEIMT'04)*, Toronto, Canada, 2004. Available online at http://www.tinwisle.com/modeling/MRS_ICEIMT04.pdf (2.10.2006).
- [McFadden, 1996] Fred R. McFadden, Data warehouse for EIS: some issues and impacts. In: *Proc. of the 29th Hawaii international Conference on System Sciences (HICSS '96) Volume 2: Decision Support and Knowledge-Based Systems*. IEEE Computer Society, Washington, USA, 1996.
- [Megill, 2005] Kenneth A. Megill, *Corporate Memory Records and Information Management in the Knowledge Age*, 2nd ed., K.G. Saur Verlag GmbH, München, 2005.
- [Moody and Kortink, 2000] Daniel L. Moody and Mark A. R. Kortink, From enterprise models to dimensional models: a methodology for data warehouse and data mart design. In: *Proc. of the Second Intl. Workshop on Design and Management of Data Warehouses (DMDW'00)*, CEUR Workshop Proceedings **28** (2000), CEUR-WS.org, 5.1 - 5.12.
- [Moody and Shanks, 1994] Daniel L. Moody and Graeme G. Shanks, What makes a good data model? Evaluating the quality of entity relationship models. In: *Proc. of the 13th International Conference on the Entity-Relationship Approach (ER '94)*, Lect. Notes Comput. Sci., **881** (1994), Springer-Verlag, London, 94-111.
- [Moss and Atre, 2003] Larissa Terpeluk Moss and Shaku Atre, *Business Intelligence Roadmap: the Complete Project Lifecycle for Decision-Support Applications*. Addison-Wesley, Boston, MA, USA, 2003.
- [Nagarajan *et al.*, 1999] Rajaram Nagarajan, Larry Whitman and S. Hossein Cheraghi, Enterprise integration. In: *Proc. of the 4th Annual International Conference on Industrial Engineering Theory, Application and Practice* (Nov. 17-20), San Antonio, Texas, 1999.

- [Navathe, 1992] Shamkant B. Navathe, Evolution of data modeling for databases. In: *Commun. ACM* **35**, 9 (Sep. 1992), ACM Press, New York, NY, USA, 112-123.
- [Oracle, 2006] Oracle Data Hubs, Oracle Corporation.
http://www.oracle.com/data_hub/index.html. (7.5.2006)
- [Oxford, 2004] A Dictionary of Computing. Oxford University Press, 2004. Oxford Reference Online. Oxford University Press. Available as <http://www.oxfordreference.com> (17.4.2006).
- [Paim *et al.*, 2002] Fábio Rilston Silva Paim, Ana Elizabete Carvalho and Jaelson Brelaz de Castro, Towards a methodology for requirements analysis of data warehouse systems. In: *Proc. of the XVI Simpósio Brasileiro de Engenharia de Software (SBES2002)*, Gramado, Rio Grande do Sul, Brazil, 2002.
- [Palvia *et al.*, 2003] Prashant Palvia, En Mao, A. F. Salam and Khalid S. Soliman, Management information systems research: what's there in a methodology? In: *Communications of the AIS* **11**, 16 (Feb. 2003), 289-309.
- [Paré, 2001] Guy Paré, Using a positivist case study methodology to build and test theories in information systems: illustrations from four exemplary studies. Cahier du GReSI #01-09 (May 2001), HEC Montréal. Also available as <http://gresi.hec.ca/SHAPS/cp/gescah/formajout/ajout/test/uploaded/cahier0109.pdf> (15.4.2007).
- [Pereira and Becker, 2000] Walter Pereira, Karin Becker, A methodology targeted at the insertion of data warehouse technology in corporations. In: K. Becker, A. Augusto de Souza, and D. Yluska de Souza Fernandes (edt.), *Proc. of the 15th Simposio Brasileiro de Banco de Dados (SBBD'00)*, October 2000. CEFET-PB Publishers, 316–330. Also available as: http://www.lbd.dcc.ufmg.br/sbbd2000/becker_pereira.pdf
- [Ponniah, 2001] Paulraj Ponniah, *Data Warehousing Fundamentals*, 1st. ed. John Wiley & Sons, Inc., New York, NY, USA, 2001.
- [Rizzi *et al.*, 2006] Stefano Rizzi, Alberto Abelló, Jens Lechtenbörger and Juan Trujillo, Research in data warehouse modeling and design: dead or alive? In: *Proc. of the 9th ACM International Workshop on Data Warehousing and OLAP (DOLAP'06)*. ACM Press, New York, NY, 2006, 3-10.
- [Sanders, 2005] Paul Sanders, Introduction to the Unified Dimensional Model (UDM). Microsoft Corporation, 2005.
<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnsq190/html/IntrotoUDM.asp>. (7.5.2006)
- [Schneider, 2003] Michel Schneider, Well-formed data warehouse structures. In: *Proc. of the 5th Intl. Workshop DMDW'2003*, CEUR Workshop Proceedings, **77** (Sep. 8, 2003). CEUR-WS.org.

- [Shanks and Darke, 1999] Graeme Shanks and Peta Darke, Understanding corporate data models. *Inform. Manag.* **35**, 1 (Jan. 1999), 19-30. Elsevier Science Publishers B. V..
- [Shin, 2002] Bongsik Shin, A case of data warehousing project management. *Inform. Manag.* **39**, 7 (Jul. 2002), Elsevier Science Publishers B. V., 581-592.
- [Shoshani, 1997] Arie Shoshani, OLAP and statistical databases: similarities and differences. In: *Proc. of the Sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS '97)*. ACM Press, New York, NY, 1997, 185-196.
- [Silverston, 2001] Len Silverston, Inmon W.H. & Graziano, K. 2001. *The Data Model Resource Book: a Library of Universal Data Models For All Enterprises*, Revised Edition, Volume 1. J. Wiley Computer Publishing, New York, 2001.
- [Sowa and Zachman, 1992] John F. Sowa and John A. Zachman, Extending and formalizing the framework for information systems architecture. In: *IBM Syst. J.* **31**, 3 (Jun. 1992). IBM Corp, USA, 590-616.
- [Staudt *et al.*, 1999a] Martin Staudt, Anca Vaduva and Thomas Vetterli, Metadata management and data warehousing. University of Zurich, Dept. of Information Technology, Technical Report 99.04, April 1999. Also available as <ftp://ftp.ifl.unizh.ch/pub/techreports/TR-99/ifi-99.04.pdf>.
- [Staudt *et al.*, 1999b] Martin Staudt, Anca Vaduva and Thomas Vetterli, The role of metadata for data warehousing. University of Zurich, Dept. of Information Technology, Technical Report 99.06, September 1999. Also available as <ftp://ftp.ifl.unizh.ch/pub/techreports/TR-99/ifi-99.06.ps.gz>.
- [Stefanov, 2006] Veronika Stefanov, Bridging the gap between data warehouses and organizations. In: *Proc. of Workshops and Doctoral Consortium, 18th Conference on Advanced Information System Engineering (CAiSE'06)*, June 2006, Namur University Press, 1160-1167.
- [Stefanov and List, 2006] Veronika Stefanov and Beate List, Business metadata for the data warehouse - weaving enterprise goals and multidimensional models. In: *Proc. of the International Workshop on Models for Enterprise Computing (IWMEC) at the 9th International Enterprise Distributed Object Computing Conference (EDOC 2006)*, Oct. 2006, 53-61.
- [Stefanov and List, 2007] Veronika Stefanov and Beate List, Explaining data warehouse data to business users - a model-based approach to business metadata. Accepted to: 15th European Conference on Information Systems (ECIS 2007), June 2007. Available online as http://www.wit.at/people/stefanov/documents/ecis07_stefanov_list_camera-ready.pdf

- [Stefanov *et al.*, 2005] Veronika Stefanov, Beate List and Josef Schiefer, Bridging the gap between data warehouses and business processes - a business intelligence perspective for event-driven process chains. In: Proc. of the 9th International Enterprise Distributed Object Computing Conference (EDOC 2005), Sep 2005. IEEE Computer Society, 3-14.
- [Thomsen, 1997] Erik Thomsen, *OLAP Solutions: Building Multidimensional Information Systems*. John Wiley & Sons, Inc, New York, NY, USA, 1997.
- [Trisolini *et al.*, 1999] Stefano M. Trisolini, Maurizio Lenzerini, Daniele Nardi, Data integration and warehousing in Telecom Italia. In: *Proc. of the 1999 ACM SIGMOD international Conference on Management of Data (SIGMOD '99)*. ACM Press, New York, NY, USA, 538-539.
- [Vassiliadis, 2000] Panos Vassiliadis, Gulliver in the land of data warehousing: practical experiences and observations of a researcher. In: *Proc. of the Second Intl. Workshop on Design and Management of Data Warehouses (DMDW 2000)*, CEUR Workshop Proceedings, **28** (2000), CEUR-WS.org 12.1-12.16.
- [Vassiliadis *et al.*, 2000] Panos Vassiliadis, Mokrane Bouzeghoub and Christoph Quix, Towards quality-oriented data warehouse usage and evolution. *Inform. Syst.* **25**, 2 (Apr. 2000), Elsevier Science Ltd, 89-115.
- [Vassiliadis and Sellis, 1999] Panos Vassiliadis and Timos Sellis, A survey of logical models for OLAP databases. *SIGMOD Rec.* **28**, 4 (Dec. 1999). ACM Press, New York, NY, USA, 64-69.
- [Watson and Gray, 1998] Hugh J. Watson and Paul Gray, *Decision Support in the Data Warehouse*. Prentice Hall Professional Technical Reference, New Jersey, USA, 1998.
- [Watson and Haley, 1997] Hugh J. Watson and Barbara J. Haley, Data warehousing: a framework and survey of current practices. *Journal of Data Warehousing* **2**, 1 (Jan. 1997), 10-17.
- [Watson *et al.*, 2002] Hugh J. Watson, Dale L. Goodhue and Barbara H. Wixom, The benefits of data warehousing: why some organizations realize exceptional payoffs. *Inform. Manag.* **39**, 6 (May. 2002), 491-502. Elsevier Science Publishers B. V.
- [Winter, 2001] Robert Winter, The current and future role of data warehousing in corporate application architecture. In: *Proc. of the 34th Hawaii International Conference on System Sciences*, IEEE, 2001.
- [Wixom and Watson, 2001] Barbara H. Wixom and Hugh J. Watson, An empirical investigation of the factors affecting data warehousing success, *MIS Quarterly* **25**, 1, (Mar. 2001), 17-41.
- [Wu and Buchmann, 1997] Ming-Chuan Wu and Alejandro P. Buchmann, Research issues in data warehousing. In: *Proc. of 7th Datenbanksysteme in Büro* (eds. Klaus

R. Dittrich and Andreas Geppert), Technik und Wissenschaft (BTW), Informatik Aktuell (1997), Springer, 61-82.

[Yin, 1994] Robert K. Yin, *Case Study Research: Design and Methods*. 2nd ed. In: Applied Social Research Methods Series, **5**. SAGE Publications Inc., Thousand Oaks (CA), London, New Delhi, 1994.

[Zachman, 1987] John A. Zachman, A framework for information systems architecture. In: *IBM Syst. J.* **26**, 3 (Sep. 1987). IBM Corp., USA, 276-292.

Liitteet

Liite 1. Kyselylomake tapaukselle 2

Huom.: kyselylomakkeen ulkoasu ei vastaa alkuperäistä kyselylomaketta, johtuen tutkielman ulkoasusta.

Ohjeet:

Vastaa omin sanoin, kuvaillen ja kriittisesti arvioiden.

Mikäli kysymyksen asettelu ei osu tehtäviisi tai toimenkuvaasi niin pyri silti vastaamaan arvioimalla projektiryhmäsi (muiden osallisten) toimintatapoja.

Vastaukset voit palauttaa joko elektronisessa muodossa (doc, pdf, jne.) tai paperiversiona.

KIITOS!

TAUSTAKYSYMYKSET

<p>Projektin tehtäväalueesi? (esim. projektinhallinta, raportointi)</p>	
<p>Työtehtäväsi ja toimenkuvasi projektissa? (esim. tietolähteiden määrittely, tietokannan fyysinen suunnittelu)</p>	
<p>Pääasialliset asiakkaan liiketoiminta-alueet ja asiakokonaisuudet, joiden kanssa työskentelit? (esim. laskutus, tavaraliikenne)</p>	
<p>Projektivastuusi? (esim. tavaraliikenteen määr. suun. ja toteutus)</p>	

Kysymykset ovat jaettu kuuteen aiheeseen ja yhdeksään kysymysosoon.

KYSYMYKSET – PROJEKTIN LAATUTAVOITTEET

1.) Määritettiinkö projektin aluksi laatutavoitteet ja miten ne saavutettaisiin?

Jos ei, niin miksi ei?

Jos määritettiin, niin minkälaisia asioita otettiin huomioon ja miten niihin pyrittiin ja miten valvottiin niiden toteutumista?

Minkälaisia asioita olisi pitänyt ottaa huomioon? (Esim. liittyen toiminnallisuuteen, käytettävyyteen, ylläpidettävyyteen, ...)

KYSYMYKSET – TIETOARKKITEHTUURI

2.) Miten projektissa huomioitiin tietojen integrointi ja organisaation yleinen tietoarkkitehtuuri?

Mihin olisi pitänyt kiinnittää erityistä huomiota? (Esim. eri lähtöjärjestelmät, organisaation yleisen arkkitehtuurin tavoitetila, koordinointi, tietovarastototeutuksen arkkitehtuuriratkaisu, ...)

KYSYMYKSET – KEHITYSPROSESSI

3.) Kuvaile käyttämäsi kehitysprosessia.

Oliko projektissa käytetty kehitysprosessi sovelias tähän projektiin?

Miten sitä olisi voinut parantaa (pl. yleinen projektinhallinta, joka ei kuulu tutkimuksen sisältöön)? (Esim. suunnitteleminen, määrittely, kehittäminen, hallinta, ... vertaa esim. johonkin tuntemaasi ohjelmistoprosessiin)

KYSYMYKSET – TIETOVARASTON MÄÄRITTELY JA KEHITTÄMINEN

4.) Kuvaile projektissa käyttämäsi/käytettyä tietovaraston

määrittelytapaa. (Esim. apuajatuksia: [jäsentävä lähestymistapa = liiketoiminnan tavoitteista konkretiaan, osittava lähestymistapa = nykytilan analysointi esim. prosessit ja tietolähteet.] /

[tavoiteohjattu = aloitetaan organisaation tavoitteista, tieto-ohjattu = aloitetaan olemassa olevasta tiedosta, vaatimusohjattu = aloitetaan liiketoimintaihmissen kysymyksistä], ...)

Toimiko se toivotulla tavalla tai miten kehittäminen olisi kannattanut toteuttaa?

5.) Miten projektissa käytettyä tietovarastointikehystä (=määrittelee mm.

kuinka projekti hoidetaan, minkälaisia osa-alueita ja vaiheita siihen kuuluu) **olisi voinut jollakin yksittäisellä toimenpiteellä tai tavalla parantaa tai hallita paremmin?**

(Esim. suunnittelumenetelmien käyttäminen, erilaisten prosessien hyödyntäminen, menetelmät, työjärjestykset, ei-tekniset toimenpiteet...)

KYSYMYKSET – TIETO- JA SUUNNITTELUMALLIT

6.) Kuvaile projektissa käytettyjä ja käyttämiäsi tietomalleja ja

tietovaraston suunnittelu- ja mallintamistapaa. (Esim. käsitekartoitusta, käsittemallin luomista, käsitteellistä mallintamista ER-mallia hyödyntäen, dimensionaalista mallintamista...)

Mitä olisi voinut tehdä toisin tai paremmin?

KYSYMYKSET – METATIEDOT, STANDARDIT JA OHJEET

7.) Kuvaile projektissa kerättyjä metatietoja (=kuvailutieto, tietoa tiedosta, sisällöstä tai dokumentin osista, tieto tietovarastointiympäristöstä ja sen kehittämisestä, komponenteista ja prosesseista).

Miten kerättyjä tietoja hallittiin?

(Esim. liiketoimintatermejä, liiketoimintaprosesseja, tiettyjä käsitteitä, tietolähteitä, organisaatioyksiköitä, yleinen sanakirja, ...)

Mitä metatietoja olisi lisäksi pitänyt kerätä ja miksi?

(Esim. liiketoimintatermejä, liiketoimintaprosesseja, tiettyjä käsitteitä, ...)

8.) Kuvaile projektissa käyttämäsi tapaa metatietojen hallintaan.

Mitä olisi pitänyt tehdä toisin? (Esim. miten ja missä projektin vaiheessa keräsit ja dokumentoit (ja mahdollisesti mitä))

9.) Miten arvioisit projektissa käytettyjä ohjeistuksia, standardeja ja nimeämiskäytäntöjä? (Esim. minkälaisia lisätarpeita olisi ollut tai minkälaisiin asioihin olisi erityisesti pitänyt keskittyä?)

VAPAA SANA JA PALAUTE (VAPAAEHTOINEN)**Esimerkiksi:**

Ideoita, ajatuksia, kritiikkiä, parannusehdotuksia tietovarastointiprojektistasi, tietovarastointiprosessista, tietovarastointikehyksestä jne...

Yleisiä kommentteja tietovarastointihankkeista.

Palautetta kysymyksistä / tutkimuksesta.

KIITOS!

Tervetuloa hakemaan pieni ”palkkio” vastauksista!

Vastaukset ovat luottamuksellisia ja niitä käytetään vain oheisinformaatiossa (saatekirjeessä) kuvattuun tarkoitukseen. Alkuperäisiä vastauksia säilytetään vain pro gradu-opinnäytteen hyväksymiseen asti, jonka jälkeen ne saatetaan muuntaa toisenlaiseen elektroniseen muotoon.

Vastauksista voidaan poimia yksittäisiä lauseita, mutta pääasiassa vastaukset analysoidaan ja niitä saatetaan koostaa ja yleistää.

HUOM: vastauksien toisenlaisia elektronisia muotoja saatetaan joutua näyttämään työn tarkastajille ja muille kiinnostuneille tutkijoille.

Tampereella 29.3.2007

Tommi Rantanen

Tampereen Yliopisto

Tietojenkäsittelytieteiden laitos

Työn ohjaaja: Timo Niemi, FT, Yliassistentti