

**VAASAN YLIOPISTO**  
**TEKNILLINEN TIEDEKUNTA**  
**TIETOTEKNIikka**

Jarkko Lehto

**TIETOVARASTOINTI JA BUSINESS INTELLIGENCE**

Tietotekniikan  
pro gradu –tutkielma

**VAASA 2016**

## SISÄLLYSLUETTELO

LYHENTEET .....	4
KUVAT .....	5
TAULUKOT .....	6
1 JOHDANTO .....	9
2 TIETOVARASTOT .....	12
2.1 Tietovarastotyypit .....	13
2.2 Tietovarastoarkkitehtuurit .....	15
2.3 Tietokannat.....	17
2.4 Tietokannan rakenne .....	18
2.5 Tietovarastoprojekti .....	19
2.6 Tietovaraston palvelun tarjoajat.....	24
3 TIEDON MUOKKAAMINEN .....	26
3.1 ETL-prosessi .....	26
3.2 ETL-työkalut .....	27
3.3 Tiedon louhinta .....	28
3.4 Big Data .....	30
3.5 Metadata.....	31
3.6 Tietovarastoinnin apuvälineet .....	32
4 BUSINESS INTELLIGENCE .....	34
4.1 Business Intelligence pilvessä.....	36
4.2 Business Intelligence sovellukset.....	37
4.3 Päätöksenteon apuvälineet .....	38
4.4 Raportointi.....	41
4.5 Mittaristot.....	43
4.6 BI-projektin toteutus .....	44

5 LIIKETOIMINNAN SUORITUSKYVYN JOHTAMINEN .....	47
5.1 Liiketoiminnan suorituskyvyn johtamisen eroaminen BI:stä .....	47
5.2 Strategia.....	48
5.3 Strategisen johtamisen edut.....	50
5.4 Strategia tasot .....	54
5.5 Suorituskykymittarit.....	55
5.6 Balanced scorecard.....	56
5.7 Six Sigma .....	57
6 YHTEENVETO .....	59
LÄHDELUETTELO .....	61

## LYHENTEET

BAM	Business Activity Monitoring
BI	Business Intelligence
BIaaS	Business Intelligence as a Service
BSC	Balanced Scorecard
CRISP	Cross Industry Standard Process for Data Mining
DSS	Decision Support System
DW	Data Warehouse
EDW	Enterprise Data Warehouse
EII	Enterprise Information Integration
EIS	Executive Information System
ETL	Extract, Transform, Load
KPI	Key Performance Indicator
LDP	Linked Data Platform
LOD	Linked Open Data
MDM	Master Data Management
OLAP	Online Analytical Process
OLTP	Online Transaction Process
RTDW	Real Time Data Warehouse
SQL	Structured Query Language
VDW	Virtual Data Warehouse

## KUVAT

Kuva 1. Tietovarastoon yhdenmukaistetaan eri tietolähteiden tietoja (Hovi, Hervonen & Koistinen 2009: 23).

Kuva 2. Itsenäiset datamartit (Hovi ym. 2009:26).

Kuva 3. EDW-arkkitehtuuri (Hovi ym 2009: 27).

Kuva 4. Kolmitasoarkkitehtuuri (Turban ym.2011: 58).

Kuva 5. Web-pohjainen tietovarasto (Turban ym. 2011: 59).

Kuva 6. Tähtimalli (Hovi, Huotari & Lahdenmäki 2005: 136).

Kuva 7. Spiraalimalli (Hovi ym. 2009: 133).

Kuva 8. ETL-prosessi (Turban ym. 2011: 68).

Kuva 9. Tiedon louhinnan kuusi vaihetta (Turban ym. 2014: 207).

Kuva 10. Business Intelligence silmukka (Giovinazzo 2000: 2).

Kuva 11. Päätöksenteon järjestelmä (Turban ym. 2014: 85).

Kuva 12. OLAP-kuutio. Nolla ulottuvuutta (Hovi ym. 2009: 91).

Kuva 13. Ylempänä yksiulotteinen ja alempana kaksiulotteinen OLAP-taulu (Hovi ym. 2009: 91).

Kuva 14. Kolmiulotteinen OLAP-kuutio (Hovi 2009:92).

Kuva 15. Tietovarastonkehitys suunnitelma (Hovi ym. 2009: 180)

Kuva 16. Strategisen suunnittelun tuomat edut (David ym:49).

Kuva 17. Strategiatasoista vastuussa olevat henkilöt (David ym. 2015: 137).

## TAULUKOT

Taulukko 1. EDW ja Data Martti lähestymistapojen erot (Turban ym 2014:311).

Taulukko 2. Tietovarastotuotteiden palveluntarjoajat (Turban ym. 2014:310).

Taulukko 3. ETL-työkalut (Hovi ym. 2009: 60).

Taulukko 4. Metadata tyyppejä (Giovinazzo 2000: 44).

Taulukko 5. Business Intelligence-työkalut (Turban ym. 2011: 41).

Taulukko 6. Seitsemäntoista ohjetta tehokkaaseen strategiaan (David ym. 2015:52).

Taulukko 7. Strategiakartta ja tasapainotettu tuloskortti (Turban ym. 2014: 361).

## TIIVISTELMÄ

---

### VAASAN YLIOPISTO

#### Teknillinen tiedekunta

<b>Tekijä:</b>	Jarkko Lehto	
<b>Tutkielman nimi:</b>	Tietovarastointi ja Business Intelligence	
<b>Ohjaajan nimi:</b>	Tero Vartiainen	
<b>Tutkinto:</b>	Kauppätieteiden maisteri	
<b>Oppiaine:</b>	Tietotekniikka	
<b>Opintojen aloitusvuosi:</b>	2007	
<b>Tutkielman valmistumisvuosi:</b>	2016	<b>Sivumäärä:</b> 62

---

### TIIVISTELMÄ:

Tutkielmassa tutkitaan tietovarastoinnin ja Business Intelligencen tekniikoita ja käytäntöjä. Tutkimus suoritettiin kirjallisuustutkimuksena ja siinä perehdyttiin tietovarastotyyppeihin ja arkkitehtuureihin, tietokantoihin, ETL-työkaluihin sekä meta-dataan. Tutkimuksessa tutkittiin myös Business Intelligenceä ja muita päätöksenteon apuvälineitä sekä raportointia ja mittaristoja. Strategista johtamista esitellään myös lyhyesti, koska Business Intelligence ja siitä saadun tiedon perustella tehdään strategisia päätöksiä ja ohjataan yritystä menestykseen.

Tutkimukset osoittivat, että tietovarastointi ja Business Intelligence voidaan toteuttaa monella eri tavalla ja monessa eri mittakaavassa. Pienimmillään ratkaisut voidaan toteuttaa pelkän excelin avulla ja laajimmillaan koko organisaation kattavalla tietovarasto ja Business Intelligence -ratkaisulla. Business Intelligence -ratkaisut toteutetaan yleensä yhdessä tietovaraston kanssa ja kustannussyistä molemmat ratkaisut ostetaan samalta valmistajalta tai toimittajalta. Business Intelligencen puuttumisesta on tänä päivänä yritykselle pelkästään haittaa ja hyvin toteutettuna se on yritykselle merkittävä kilpailuetu. Parhaimmillaan Business Intelligence antaa yrityksen johdolle reaaliaikaista tietoa päätöksenteon tueksi.

---

**ASIASANAT:** Data Warehousing, Business Intelligence, Tietovarastointi

## ABSTRACT

---

### UNIVERSITY OF VAASA

Faculty of technology

**Author:**

Jarkko Lehto

**Topic of the Master's Thesis:**

Data warehousing and Business Intelligence systems

**Instructor:**

Tero Vartiainen

**Degree:**

Master of Science in Economics and Business Administration

**Major:**

Computer Science

**Year of entering the University:**

2007

**Year of completing the Master's Thesis:** 2016

**Pages:** 62

---

### ABSTRACT:

This study examines techniques and practices of data warehousing and business intelligence. Study was conducted as literature study and it familiarizes the types of data warehouses, architectures, databases, ETL-tools and meta-data. In the study was also included other decision support systems, reporting and reports. Strategic management was studied shortly in this study as well.

The studies showed that business intelligence can be implemented many different ways and many different scales. The smallest solution could be just Microsoft Excel report. Largest scale of business intelligence is organization wide data warehouse and business intelligence system. Now a day's business intelligence system is essential to all companies to succeed on the market. However, business intelligence does not guarantee success on the market. It gives the needed information to executives to make better strategic decisions.

---

**KEYWORDS:** Data Warehousing, Business Intelligence

## 1 JOHDANTO

Yritysten liiketoimintaympäristö muuttuu ja elää jatkuvasti. Ympäristön muuttuminen vaatii yritysten johdolta enemmän ja nopeampia päätöksiä pärjätäkseen markkinoilla. Liiketoiminta ympäristön muuttuminen monimutkaisemmaksi antaa yritykselle mahdollisuuksia, mutta toisaalta aiheuttaa taas ongelmia (Turban, Sharda & Delen 2014: 5). Päätöksen tekoa varten yrityksen johto tarvitsee tietoa menneestä, jotta voisi tehdä oikeita ratkaisuja saadakseen yrityksen menestymään. Tietoa kerätään yritysten operatiivisista järjestelmistä ja tytäryhtiöiden järjestelmistä. Lisäksi kerätään ulkoista tietoa esimerkiksi kilpailijoista sekä kilpailuympäristöstä. Ilman tietovarastointia ja Business Intelligenceä sekä niiden tuomaa tiedon prosessointitehoa oikeita päätöksiä ei voitaisi tehdä riittävän nopeasti.

Tutkielmassa perehdytään tietovarastoinnin ja Business Intelligencen perusasioihin. Tietovarastoista esitellään erilaiset tietovarastotyytit, -arkkitehtuurit sekä hieman tietokannan rakennetta. Koska operatiivisesta järjestelmästä ei voi suoraan tallentaa tietoa tietovarastoon, perehdytään myös tiedon muuntamiseen tietojärjestelmään sopivaksi. Tiedon tallentamisen lisäksi tutkielmassa kerrotaan mitä metadata on ja miten sitä voidaan hyödyntää tietovarastoissa.

Business Intelligence on varsin uusi käsite tiedon hyödyntämisessä. Tutkielmassa kerrotaan Business Intelligencen perusteet sekä hieman sen historiasta. Business Intelligencen ollessa kuitenkin laaja käsite sekä tulkittu hieman monikäsitteisesti, tutkielmassa perehdytään hieman muihinkin päätöksenteon apuvälineisiin. Raportointiin sekä mittaristoihin perehdytään myös hieman.

Tutkielmassa halutaan kuvata tietovarasto ja Business Intelligence tekniikoita ja käytäntöjä yleisellä tasolla, eli tässä tekstissä ei perehdytä tietovarasto tai Business Intelligencen käyttöönottoprojektin toteutukseen. Käyttöönottoprojekti vaatisi valitsemaan jonkin tietyn palveluntarjoajan järjestelmän ja kuvaamaan sitä. Pelkästään toisen prosessin kuvaaminen yksityiskohtaisesti laajentaisi tutkimuksen laajuudeltaan huomattavan suureksi. Tietovarasto- ja BI-projektien toteutuksesta kerrotaan kuitenkin yleisellä tasolla.

Tutkielma ei ole syvälle luotaava tutkimus mistään tietystä tietovarasto- tai Business Intelligence ratkaisusta, vaan on aihetta yleisesti esittelevä kirjallisuustutkimus. Tutkielmassa mainitaan myös yleisimmät tietovarastointi ja Business Intelligence -

ratkaisuja tarjoavat yritykset. Tutkimuksessa ei ole perehdytty mihinkään yksittäiseen ratkaisuja tarjoavaan työkaluun eikä työkalujen tai eri valmistajien tarjoamiin ratkaisujen eroihin.

Business Intelligencen avulla tehdään myös strategisia päätöksiä. Tutkielmassa perehdytään liiketoiminnan suorituskyvyn johtamiseen ja sitä kautta yrityksen strategiaan päätöksiin. Strategiaa ei avata syvällisesti, tutkielmassa lähinnä kuvataan miten strategisia päätöksiä tehdään Business Intelligencen avulla.

Tutkimus tehdään kirjallisuuskatsauksena. Tutkimuksessa perehdytään tietovarastointiin ja Business Intelligenceen liittyvään kirjallisuuteen ja artikkeleihin. Aineistosta koostetaan kattava kuvaus tietovarastoinnista ja siihen liittyvistä asioista. Näitä asioita ovat mm. tietokannat, järjestelmien arkkitehtuurit sekä käytettävät ohjelmistot. Tutkimuksessa pyritään vastaamaan kysymyksiin: mitä on tietovarastointi ja liiketoiminta tiedon hallinta? Miten tietoa varastoidaan ja miten sitä voidaan käyttää hyödyksi? Tutkimuksessa selvitetään millaisia järjestelmiä toimiva tietovarasto ja BI-järjestelmä tarvitsee. Teoksessa käsitellään myös muita päätöksenteon apuvälineitä ja niiden luokitteluita.

Kirjallisuuskatsaus on metodi ja tutkimustekniikka, joka tutkii tutkittua tutkimusta. Kirjallisuuskatsauksella on useita tavoitteita, se voi vahvistaa olemassa olevaa teoriaa tai luoda uutta. Se voi arvioida teoriaa ja tarkistaa ovatko tiedot totuudenmukaisia ja luotettavia. Se voi myös rakentaa kokonaiskuvaa tutkimuksen kohteesta. Tämä tutkimus toteutettiin kuvailevana kirjallisuuskatsauksena. Kuvailevalle kirjallisuuskatsaukselle tyypillistä on laajat aineistot eikä aineistoa rajata metodisten sääntöjen perusteella. Narratiivinen kirjallisuuskatsaus kuvailee tutkittavan aiheen historiaa ja kehitystä sekä antaa laajan kuvan käsiteltävästä aiheesta. Se yhdistelee tutkittua tietoa eri lähteistä ja esittää sen järkevässä ja helppolukuisessa muodossa. Tässä tutkimuksessa kerätään tietoa Tietovarastoinnista, Business Intelligencestä sekä johtamisesta ja esitellään se lukijalle selkeästi johdonmukaisessa järjestyksessä. (Eskelinen & Karsikas 2014: 89, Salminen 2011: 3-7)

Tutkimuskysymyksiä on useita, ensimmäisenä kysymyksenä on ”Mitä on tietovarastointi?”. Toisena kysymyksenä on ”Mitä on Business Intelligence?”. Nämä kaksi ovat varsinaisia tutkimuskysymyksiä, mutta tutkimuksessa pyritään myös vastaamaan kysymyksen miten tietovarastointi ja Business Intelligence liittyvät johtamiseen.

Tutkielma alkaa johdannolla, jonka jälkeen esitellään tietovarasto ja siihen liittyvät arkkitehtuurit. Tietokannat ja niiden rakenne kuvataan tässä samassa yhteydessä. Tiedon

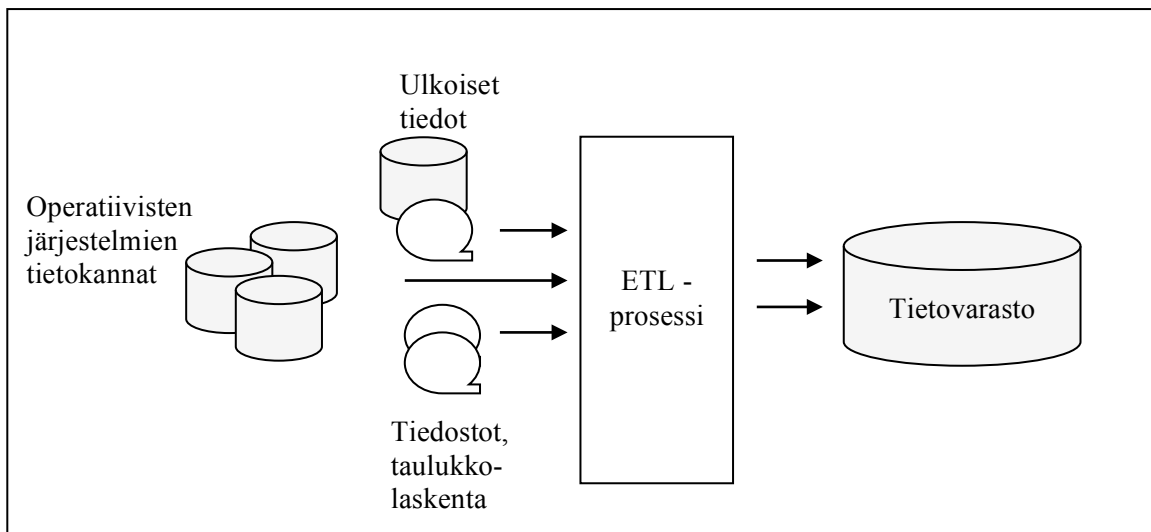
prosessointia ja siihen käytettäviä työkaluja esitellään luvussa kolme. Neljännessä luvussa perehdytään Business Intelligenceen sekä muihin päätöksenteon apuvälineisiin. Myös raportointiin ja mittaristoihin perehdytään tässä luvussa. Viidennessä luvussa kerrotaan suorituskyvyn johtamisesta. Tässä kappaleessa avataan johtamisen viitekehysä sekä kerrotaan niistä perusasiat.

## 2 TIETOVARASTOT

Tietovarasto (Data Warehouse, DW) on operatiivisesta järjestelmästä erillään oleva ympäristö, johon on tallennettu haluttu tieto halutussa muodossa yrityksen hyödynnettäväksi. Tietovarastoon ei tallenneta kaikkea mahdollista dataa, vaan sinne kerätään sellainen data, josta on yritykselle hyötyä.

Tietovarasto on tietokanta, joka on tehty avustamaan päätöksen tekoa. Tietovarasto on myös nykyisen ja menneen datan varasto. Tallennettu data on muokattu valmiiksi tiedon louhintaa, analysointia, kyselyitä ja raportointia varten. (Turban, Sharda, Delen & King 2011: 52.)

Tietovarastoon on koottu dataa yrityksen kaikista operatiivisista järjestelmistä, prosessi on kuvattu kuvassa 1.



Kuva 1. Tietovarastoon yhdenmukaistetaan eri tietolähteiden tietoja (Hovi, Hervonen & Koistinen 2009: 23)

Tietovaraston datasta voidaan tehdä erilaisia raportteja ja analyyseja rasittamatta yrityksen operatiivista järjestelmää. Yrityksellä voi olla myös yhtäaikaisesti käytössä useita operatiivisia järjestelmiä, joiden data on tallennettuna eri tietokantoihin. Operatiivisia järjestelmiä ovat mm. laskutus-, tilausten käsittely- ja toiminnanohjausjärjestelmä. Tie-

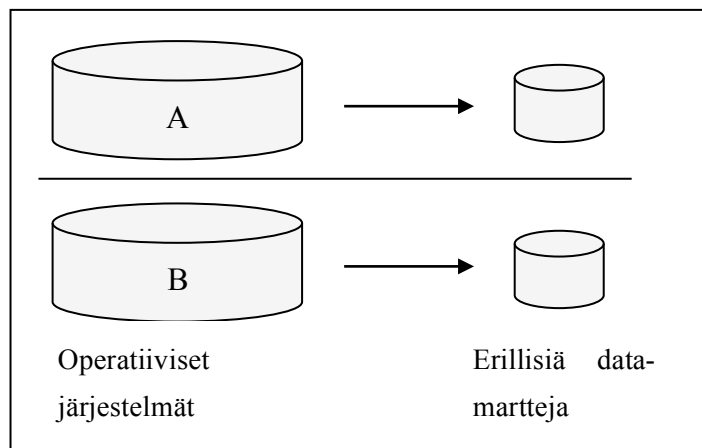
tokannat voivat olla rakenteeltaan hyvin erilaisia, joten dataa ei voida suoraan siirtää sellaisenaan tietovarastoon. Tiedon oikeellisuus tulee tarkistaa ja data täytyy muokata yhtenäiseen muotoon ennen sen siirtoa varsinaiseen tietovarastoon.

Operatiivisen tietokannan data muuttuu jatkuvasti. Esimerkiksi varastosaldo vaihtuu, kun tavara myydään. Tietovaraston tietokantaan taas jää kaikki merkinnät, mitä tietokantaan tallennetaan. Kun varastosaldo on muuttunut seuraavaan ajoon mennessä, niin uusi saldo tallennetaan, mutta vanha saldo jää myös tietokantaan. Näin nähdään varastosaldot eri ajankohtina.

## 2.1 Tietovarastotyypit

Tietovarastoja voi olla erilaisia riippuen yrityksen käyttötärpeesta. Pienempi tietovarasto on nimeltään datamartti (Hovi ym. 2009: 24). Datamartti on yksinkertainen muoto tietovarastosta, joka keskittyy yhteen aihealueeseen esim. myynti, talous tai markkinointi (kuva 2).

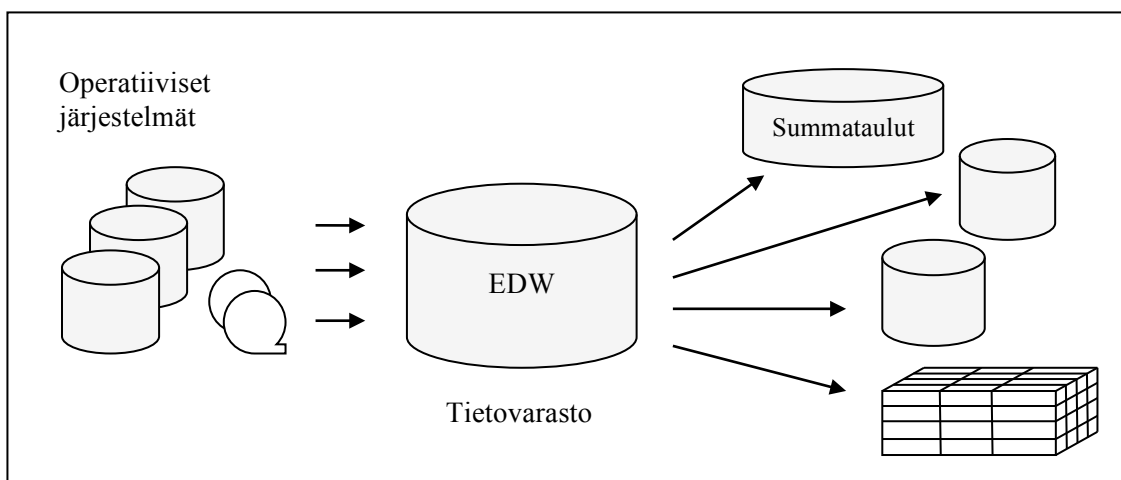
Datamartteja on kahta tyyppiä, riippuvainen ja riippumaton datamartti. Riippuvainen datamartti on koostettu tietovaraston datasta. Tietovarastoon tallennettu data on jo kertaalleen tarkistettu ja muokattu oikeaan muotoon, joten tietomalli on johdonmukainen ja data on laadukasta. Yhtenäisestä tietomallista johtuen koko organisaation raportit ovat yhtenäisiä. Riippuvainen datamartti on kallis toteuttaa, koska se vaatii tietovaraston toimiakseen. (Turban ym. 2011: 53.)



Kuva 2. Itsenäiset datamartit (Hovi ym. 2009:26).

Riippumattomat datamartit ovat muuten samanlaisia kuin riippuvaiset datamartit, mutta niiden tiedon lähteenä ei ole tietovarasto. Riippumattomiin datamartteihin kerätään usein tietoa vain yhdestä lähteestä, esim. operatiivisesta järjestelmästä. (Oracle 2013). Riippumattomat datamartit ovat suhteellisen helppoja ja nopeita toteuttaa, joten niitä tarvitseva osasto yleensä toteuttaa ja ylläpitää niitä itse. Tästä johtuen yrityksessä voi olla useita samankaltaisia datamartteja. Mikäli datamartteihin haetaan tietoa operatiivisesta tietokannasta, tiedonhaku voi hidastaa yrityksen operatiivisen järjestelmän toimintaa.

Keskitetty tietovarasto (Enterprise Data Warehouse, EDW, kuva 3) kokoaa organisaation tiedot yhteen tai muutamaa isoon tietokantaan. Tietokanta pitää sisällään usean liiketoiminta-alueen tiedot yhdenmukaistettuna. (Hovi ym. 2009: 27.) Tietovarastot on yleensä tehty yrityksen IT-osaston toimesta. Keskitetty tietovarasto palvelee koko yritystä.



Kuva 3. EDW-arkkitehtuuri (Hovi ym 2009: 27).

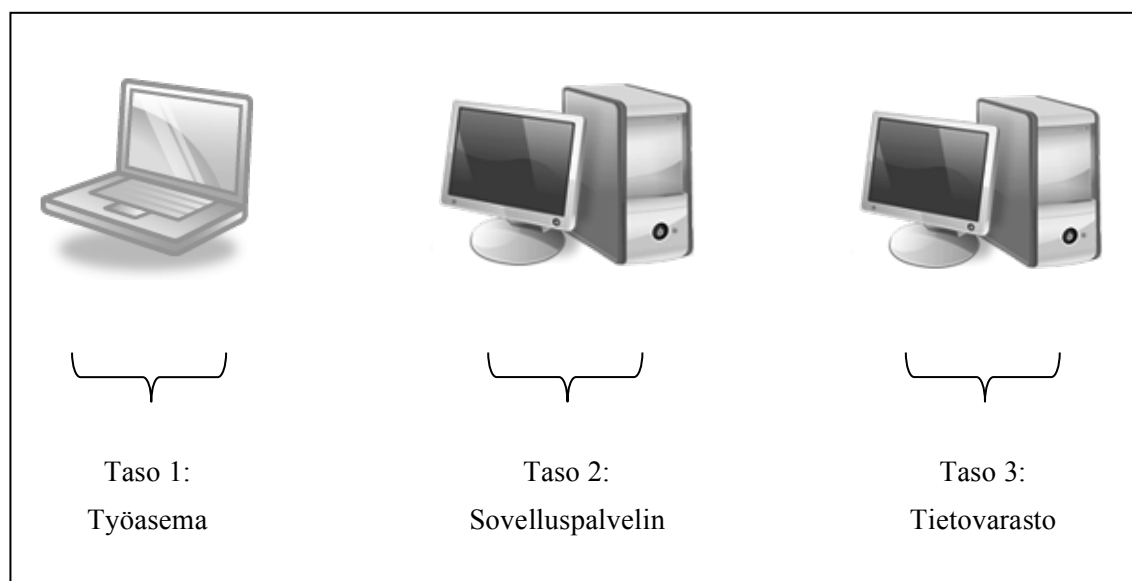
IT-osasto ylläpitää tietovarastoa ja tekee niistä valmiita summatauluja, datamartteja ja kuutioita. Loppukäyttäjät tekevät kyselyitä sitten valmiista ja helppokäyttöisistä datamarteista. Tietovarastosta muodostetut datamartit voivat pitää sisällään osaston ja järjestelmien rajojen ulkopuolista dataa. Tietovarastoista tehdyt datamartit eivät myöskään rasita yrityksen operatiivisia järjestelmiä. EDW-arkkitehtuurista on olemassa myös mal-

li, jossa ei käytetä datamartteja ja summatauluja. Tässä mallissa käyttäjät tekevät haut suoraan tietovaraston tietokannasta. Arkkitehtuurin etuna on se, että kaikki tieto on kaikkien käytettävissä. Tieto ei rajoitu valmiiden datamarttien tarjoamaan dataan.

Kolmantena arkkitehtuurimallina on yhdenmukaistetut riippumattomat datamartit. Mallissa datamartit rakennetaan tähtimallimenetelmällä yhteismitallisiksi, jolloin osa dimensioista on yhdenmukaisia. Esimerkiksi tuote tai asiakas dimensiot esiintyvät useissa eri tähtimalleissa. (Hovi ym. 2009: 28, Turban ym. 2011: 62.) Riippumattomat datamartit ovat yksinkertaisia ja helppoja toteuttaa, joten ne sopivat pienempiin yrityksiin, joilla ei ole mahdollisuuksia panostaa keskitettyyn ratkaisuun. Datamarttien laajennettavuus on kuitenkin rajallista, joten niiden käyttö on aina suunniteltu tietylle kohderyhmälle.

## 2.2 Tietovarastoarkkitehtuurit

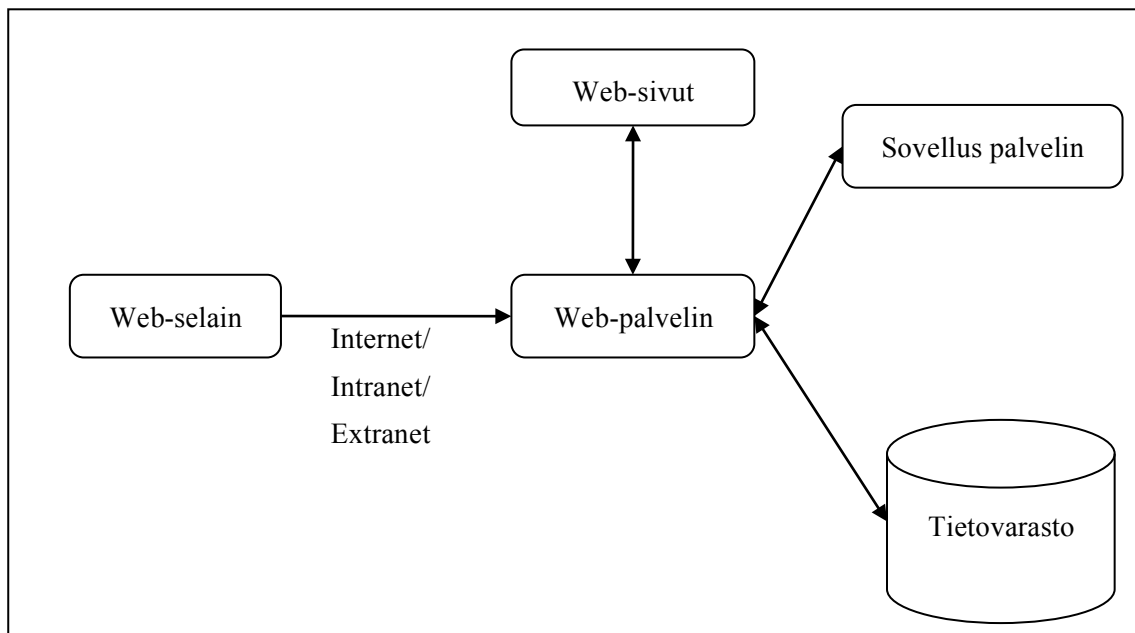
Tietovarastoarkkitehtuurina voidaan käyttää useita perinteisiä tietojärjestelmäarkkitehtuureja. Useimmiten käytetään asiakas/palvelin tai n-tason arkkitehtuuria. Kolmitasoarkkitehtuurissa tietovarasto sijaitsee erillisellä palvelimella ja ohjelmat, joilla tietoa haetaan tietovarastosta sijaitsevat toisella palvelimella. Kolmitaso arkkitehtuuri on kuvattu kuvassa 4.



Kuva 4. Kolmitasoarkkitehtuuri (Turban ym.2011:58).

Asiakas/palvelin arkkitehtuurissa (kaksitasoarkkitehtuuri) päätöksenteko- ja Business Intelligence –sovellukset sijaitsevat samalla palvelimella tietovaraston kanssa. Kaksitasoarkkitehtuuri on edullisempi toteuttaa, koska siinä tarvitsee vähemmän laitteistoa. Haittapuolena siinä voi olla suorituskykyongelmia, mikäli kyseessä on suuri tietovarasto. (Turban ym. 2011: 62.)

Web-pohjainen tietovarastointi (kuva 5) voidaan toteuttaa joko kaksi- tai kolmitasoarkkitehtuuria hyväksikäyttäen. Kolmitasoarkkitehtuurissa web-selain ottaa yhteyden web-palvelimeen, joka taas pyytää sovelluspalvelinta noutamaan tarvittavan tiedon tietovarastosta. Noudon jälkeen sovelluspalvelin palauttaa tiedot web-palvelimelle, joka näyttää ne selaimelle. Web-pohjainen tietovarasto on helppo pääsyinen, alustariippumaton ja edullinen toteuttaa.



Kuva 5. Web-pohjainen tietovarasto (Turban ym. 2011:59).

Yhtenä tyyppinä mainittakoon vielä reaaliaikainen tietovarasto (Real Time Data Warehouse, RTDW). Reaaliaikainen tietovarasto sopii tilanteisiin, joissa tarvitaan reaaliaikaista dataa päätöksenteon tueksi. Normaalin tietovaraston data ei pelkästään riitä tähän tarkoitukseen, koska se voi olla jo vanhentunutta. Tätä varten on kehitetty virtuaalinen tietovarasto (Virtual Data Warehouse, VDW). Virtuaaliseen tietovarastoon kerätään tietovarastosta historiallinen data sekä operatiivisista järjestelmistä reaaliaikainen data.

Loppukäyttäjälle virtuaalinen tietovarasto, joka siis yhdistää tietovaraston ja operatiivisten järjestelmien tiedot, näkyy yhtenäisenä raporttina, jossa on sekä historiallinen että reaaliaikainen data. Tällöin päättäjillä on käytössään tietovaraston tarjoama jalostettu data, sekä sen lisäksi operatiivisista järjestelmistä saatu reaaliaikainen tieto. (Farrah 2013)

### 2.3 Tietokannat

Yrityksellä voi olla useita erilaisia tietokantoja käytössään. Esimerkiksi fuusion yhteydessä toisella yhtiöllä on ollut Microsoftin tarjoama tietokanta ja toisella yrityksellä on ollut Oraclen tietokanta. Fuusion yhteydessä on kuitenkin päätetty jatkaa käyttäen kahta eri operatiivista järjestelmää ja siirtää uuden yhteisen järjestelmän käyttöönotto myöhemmälle. Tällaisessa tilanteessa datan Business Intelligence-käyttö vaatii runsaasti it-osaston toimenpiteitä, jotta data saadaan muokattua yhtenäiseen muotoon raportteja varten.

Operatiiviset järjestelmät tarvitsevat nopean pääsyn tietokantoihin ja vaativat useita samanaikaisia yhteyksiä tietokantoihin. Operatiiviset järjestelmät on toteutettu tapahtumankäsittelyjärjestelmiksi, OnLine Transaction Processing Systems (OLTP) (Hovi ym. 2009: 22). Operatiivisen järjestelmän tietokannan tulee olla nopea ja palvella useita käyttäjiä samanaikaisesti, kun taas tietovarastotietokantaa ei päivitetä päiväsaikaan ollenkaan.

Tietovaraston tietokanta pitää sisällään datat useilta vuosilta ja se toimii ikään kuin organisaation muistina ja mahdollistaa trendianalyysien teot (Hovi ym. 2009: 22). Tietovarastotietokannan tulee olla rakenteeltaan selkeä, helppokäyttöinen ja ymmärrettävä. Taulut ja sarakkeet tulee olla nimetty selkeästi. Tietokannasta tulee voida tehdä kyselyjä ilman varsinaista ohjelmointia (Hovi, Huotari & Lahdenmäki 2005).

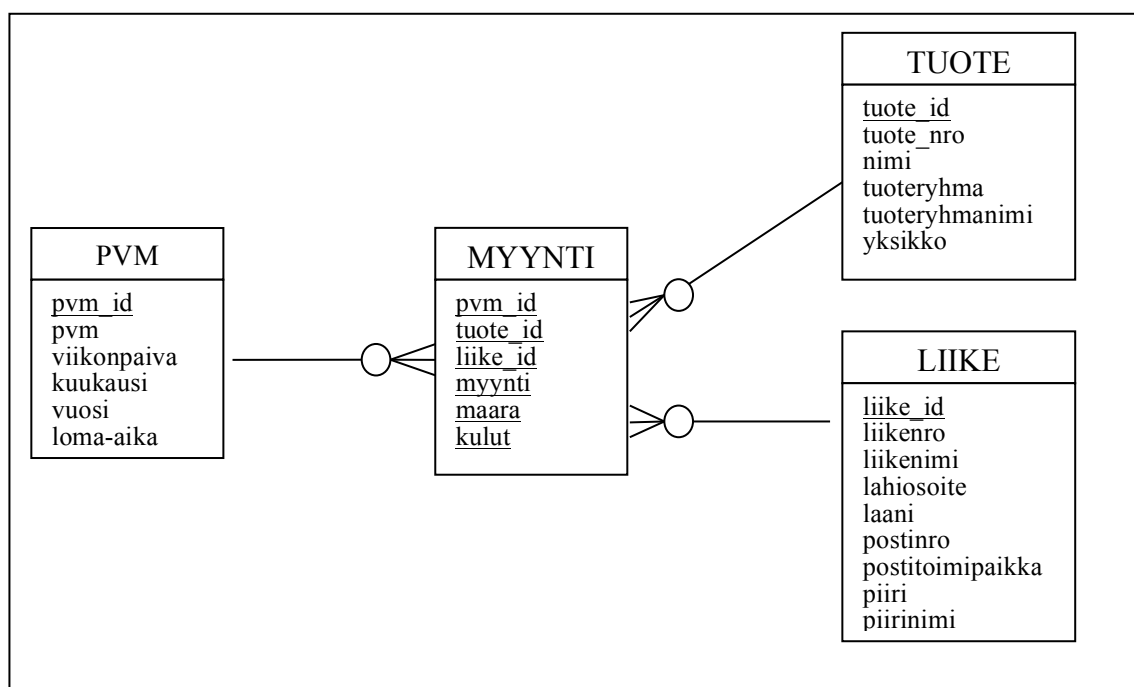
Tietovarastojen tietokantoina käytetään yleisimmin suurimpien valmistajien kaupallisia tietokantoja, kuten IBM:n DB2, Oracle ja Microsoftin SQL Server. Nämä tietokannat on alun perin suunniteltu operatiivisiin järjestelmiin ja ne sisältävät raskaita, ajantasaista tapahtumankäsittelyä tukevia ominaisuuksia. Myöhemmin tietokannan hallintajärjestelmiin on lisätty ETL- ja raportointiominaisuuksia. (Hovi ym. 2009: 62.)

## 2.4 Tietokannan rakenne

Tietovaraston tietokanta on suunniteltava selkeäksi rakenteeltaan, jotta sitä on helppo käyttää. Taulujen ja tietojen nimeämisessä on pyrittävä selkeyteen, koska tietovaraston loppukäyttäjänä on myös IT-osaston ulkopuolisia henkilöitä. (Hovi ym. 2009:36.)

Tähtimallia (Star Schema) on usein käytetty moniulotteisen ja numeerisen tiedon tallentamiseen. Tähtimallissa on yksi faktataulu ja useita dimensiotauluja. Faktataulu on normalisoitu ja dimensiotaulut ovat denormalisoidussa muodossa. Kuvassa 6 on kuvattu tähtimalli, jossa faktataulu MYYNTI on normalisoitu. Mikäli dimensiotaulutkin normalisoitaisiin, malli muuttuisi lumihiihtalemalliksi. (Hovi ym. 2005: 136.)

Tietovaraston suunnittelussa on kiinnitettävä erityistä huomiota indeksointiin. Operatiivisista järjestelmistä saatavat luonnolliset perusavaimet tulee korvata surrogaateilla, eli keinotekoisilla avaimilla. Tämä on tärkeää, koska operatiivisen järjestelmän perusavaimet voivat olla esimerkiksi tuotenumeroita, jotka voidaan ottaa uudelleenkäyttöön myöhemmin. Saman avaimen käyttö kahdessa eri tuotteessa sotkisi tietovarastosta tuotetut raportit. Surrogaattina tulisi käyttää juoksevaa numerointia. Globaalien yksilöllisten avaimien käyttöä tulisi välttää, koska ne vievät enemmän levytilaa ja ovat hitaampia käyttää. (Hovi ym. 2005:136.)



Kuva 6. Tähtimalli (Hovi ym. 2005: 136).

Kuvan 6 mukaisessa tähtimallin tietokannassa tulee eniten rivejä ”MYYNTI”-tauluun. Dimensiotauluihin tulee harvemmin lisäyksiä. Esimerkiksi myyntitapahtumia lisätään useita kertoja päivässä, mutta asiakkaita lisätään silloin tällöin. Näitä dimensiotauluja kutsutaan hitaasti muuttuviksi dimensioiksi. Tietovarastoa suunniteltaessa on tehtävä päätös, halutaanko historia tiedot pitää sellaisenaan vai muutetaanko historiaa. Esimerkiksi kun tuotenumero käytetään uudelleen ja tietovaraston dimensiotauluihin haetaan data Master data taulusta, niin silloin tuotteen historiatiedot muuttuvat. Tai itse asiassa uudella tuotteella on jo valmiiksi myyntiä. Dimensiodien muutoksen hallintaan on monia tapoja, mutta yleisimmin käytetään kolmea Ralph Kimballin tapaa:

1. Kirjoitetaan muutokset vanhan tiedon päälle.
2. Tehdään uusi tietue muuttuneilla tiedoilla ja jätetään vanha.
3. Dimensiotauluun laitetaan uusi sarake, johon tulee uusi arvo.

Mikäli dimensiotauluihin tulee muutoksia, summataulut on luotava muuttuneilta osin uudestaan. (Hovi ym. 2005:140.)

Tietovarastoa suunniteltaessa on myös tärkeää ottaa huomioon tietokantojen fyysinen suunnittelu sekä levytilan tarpeen arviointi. Fyysisessä suunnittelussa on otettava huomioon tietokantasivun koon ja tyhjän tilan määrittäminen, taulujen ja indeksien ositus, taulujen tilavaraus, sijoittelurajoitukset sekä mahdolliset tietokantapuskuriallasrajoitukset. Suunnittelijan on tunnettava hyvin tietokannanhallintajärjestelmä, jotta ominaisuudet saadaan määriteltyä huolellisesti. (Hovi ym. 2005: 142.)

## 2.5 Tietovarastoprojekti

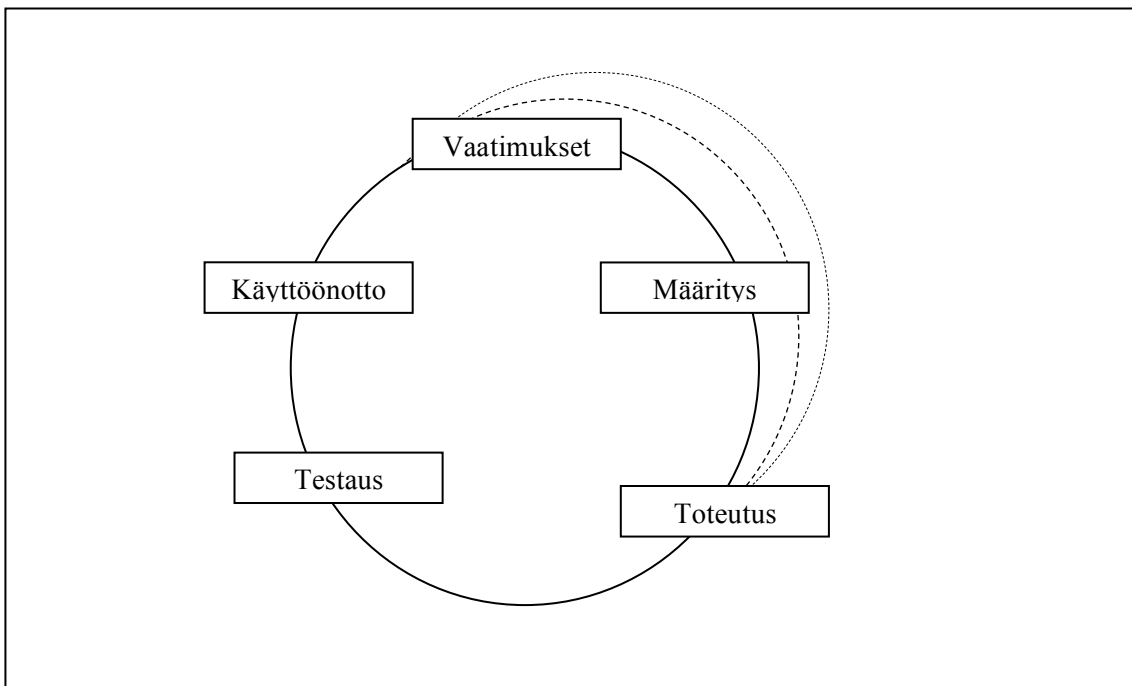
Tietovarastoprojektin toteuttamiseen pätevät samat mallit ja menetelmät kuin operatiivisen järjestelmän toteuttamiseen. Tietovarastoprojekteissa, kuten muissakin it-projekteissa, tulee muutoksia kesken prosessin. Tämän takia tietovarasto tulisi rakentaa iteraatiota hyväksikäyttäen. Tietovaraston suunnittelussa tulee huomioida käytettävissä olevat operatiiviset järjestelmät sekä tiedon laatu, joka niistä halutaan tuoda tietovarastoon. (Hovi ym. 2009: 130.)

Tietovaraston kehittämisessä on syytä käyttää menetelmiä, joilla saadaan nopeasti tuloksia. Hyväksi todettuja malleja ovat inkrementaalinen rakentaminen sekä spiraalimalli. Ensin mainitussa mallissa tietovarastoa rakennetaan inkrementti kerrallaan, esim.

ensin taloustiedot, toiseksi myyntitiedot jne. Ensimmäisen inkrementin rakentamisen on erityisen tärkeää, koska tässä vaiheessa testataan käytetyt menetelmät ja välineet.

Kun menetelmät ja välineet on todettu toimiviksi, voidaan seuraavat inkrementit toteuttaa nopeammin. Inkrementaalimalli on kuin palapeli: jokaisen inkrementin on sovittava toistensa kanssa yhteen. (Hovi ym. 2009: 132.)

Spiraalimallissa (kuva 7) ei aseteta lopullisia tavoitteita projektin alkuvaiheessa, vaan



Kuva 7. Spiraalimalli (Hovi ym. 2009: 133).

rakennetaan tietovaraston ensimmäinen versio nopeasti. Ensimmäistä versiota kommentoidaan ja siihen esitetään parannusehdotuksia, jotka toteutetaan spiraalimallin seuraavalla kierroksella. Uusia kierroksia voidaan aloittaa myös ennen edellisen kierroksen valmistumista, jolloin saadaan nopeutettua tietovaraston valmistumista. (Hovi ym. 2009: 132.)

Inkrementaalista rakentamista kutsutaan myös Data Martti lähestymistavaksi ja tätä lähestytään alhaalta-ylös mallin mukaan. Tässä prosessissa tietovarasto rakennetaan Data Martti kerrallaan. Spiraalimallia kutsutaan EDW lähestymistavaksi, jota lähestytään ylhäältä-alas mallin mukaan. Tässä tapauksessa rakennetaan ensin organisaation

kattava tietovarasto. Taulukossa 1 on kuvattu inkrementaali ja EDW lähestymistapojen erot. (Turban ym. 2014, s311-312)

Taulukko 1. EDW ja Data Martti lähestymistapojen erot (Turban ym. 2014: 311).

<b>Panostus</b>	<b>Data Martti lähestymistapa</b>	<b>EDW lähestymistapa</b>
Laajuus	Yksi kohdealue	Useita kohdealueita
Kesto	Kuukausia	Vuosia
Kustannukset	\$10 000 - \$100 000+	\$1 000 000+
Haastavuus	Pieni tai keskisuuri	Suuri
Tiedon jaettavuus	Liiketoiminta-alueen laajuinen	Organisaation laajuinen
Tietolähteet	Joitain ulkoisia järjestelmiä	Useita operatiivisia ja ulkoisia lähteitä
Koko	Megatavuista useihin gigatavuihin	Gigatavuista petatavuihin
Tiedon muunnokset	Lähiaikainen ja historiallinen data	Historiallinen data
Päivitys tiheys	Tunneittain, päivittäin, viikoittain	Viikoittain tai kuukausittain
Laitteisto	Työasemat ja osaston palvelimet	Organisaation palvelimet ja mainframe koneet
Käyttöjärjestelmät	Windows ja Linux	Unix, Z/OS, OS/390
Tietokannat	Työryhmä tai tavalliset tietokantapalvelimet	Organisaation tietokantapalvelimet
Yhtäaikaisia käyttäjiä	Kymmeniä	Sadoista tuhansiin
Käyttäjätyyppit	Liiketoiminta analyytikot ja päälliköt	Organisaation analyytikot ja johtajat
Liiketoiminta tarkoitus	Aktiviteettien optimointi liiketoiminta-alueella	Toimintojen välinen optimointi ja päätöksen teko

Tietovarastoprojekti on suuri panostus, joka täytyy suunnitella ja toteuttaa huolellisesti. Tietovarastoprojektin onnistuminen vaatii, että projektista on todellista hyötyä yritykselle. Projektilla tulee olla omistaja ja taho, joka haluaa projektin onnistuvan. Projektin laajuus tulee rajata tarkasti ja on mietittävä mitkä liiketoiminta-alueet otetaan mukaan projektiin sekä mistä operatiivisista järjestelmistä tietoa tuodaan. Projektista saatujen hyötyjen tulee olla myös mitattavissa, jolloin projektin jälkeen voidaan selvittää miten projekti on onnistunut. Tietovarastoprojektin toteuttamisessa kannattaa hyödyntää jo aiemmin todettuja parhaita käytäntöjä, joita ovat:

- Projektin tulee sopia strategiaan ja liiketoimintatavoitteisiin
- Projektilla täytyy olla johtajien, päälliköiden ja käyttäjien tuki
- Tietovarasto tulee rakentaa käyttäen inkrementaalimallia
- Projektin / tekijöiden täytyy olla sopeutuvaisia
- Projektia tulee johtaa sekä IT:n että liiketoiminnan puolelta
- Tilaaja – toimittajasuhde täytyy olla selvä
- Käytä vain puhdistettua ja laadukasta dataa
- Älä unohda koulutusta

Tietovarastoprojektissa on paljon riskejä. Näitä samoja riskejä on myös muissa IT-projekteissa, mutta tietovarastointiprojektin riskeihin täytyy suhtautua vakavammin, koska tietovarastoprojektit ovat laajoja ja kalliita. (Turban ym. 2014: 314-316, Hovi ym. 2009: 133.)

Onnistuneen tietovarastoprojektin toteutuksessa kannattaa harkita seuraavia asioita:

- **Älä aloita väärän rahoittajan kanssa.** Tarvitset johtotason tuen ja rahoituksen, jolla on vaikutusvaltaa tarvittaviin resursseihin. Tarvitset myös johtotason projektin vetäjän, joka on ansainnut muiden johtajien kunnioituksen ja jolla on terve epäily tietotekniikka kohtaan. Tarvitse myös IT-päällikön tukemaan projektia.
- **Asetat tavoitteita, joita et voi saavuttaa.** Jokaisessa tietovarastoprojektissa on kaksi vaihetta: myyntivaihe, jossa markkinoit projektin tuomia etuja ja toisessa vaiheessa taistelet saavuttaaksesi tavoitteet, jotka asetit vaiheessa 1.
- **Annat väärin lupauksia projektista.** Älä väitä, että tietovarastoprojekti auttaa johtajia tekemään parempia päätöksiä. Tietovarasto tarjoaa tarvittavan tiedon, jotta johtajat voivat tehdä parempia päätöksiä.

- **Täytät tietovaraston tiedolla vain siksi kun sitä on saatavilla.** Tietovarasto ei ole kaatopaikka, kerää vain sellaista dataa, jota tarvitaan.
- **Uskot, että tietovaraston tietokanta on rakenteeltaan samanlainen kuin operatiivisen järjestelmän tietokanta.** Yleensä se ei ole. Operatiivisen järjestelmän tietokanta on yleensä normalisoitu relaatiotietokanta, kun taas tietovaraston tietokanta ei ole normalisoitu ja on moniulotteinen.
- **Valitset tietovarastopäällikön, joka on enemmän teknologia orientoitunut kuin käyttäjä orientoitunut.** Tietovaraston päätarkoitus on toimittaa liiketoiminnalle heidän tarvitsemaa tietoa. Tietovarastoa ei tule rakentaa vain teknologian vuoksi.
- **Keskityt vain perinteiseen sisäiseen dataan ja unohdat ulkoisen datan merkityksen.** Dataa saadaan monessa muodossa myös yrityksen ulkopuolelta. Datan tulee olla oikeiden ihmisten saatavilla oikeaan aikaan oikeassa muodossa.
- **Toimitat päällekkäistä data sekavin määrittelyin.** Datan putsaaminen on erittäin tärkeä prosessi tietovarastoinnissa. Se sisältää tiedon muokkaamisen oikeaan muotoon sekä päällekkäisen datan poiston.
- **Uskot lupauksiin suorituskyvystä.** Tietovarastot vaativat yleensä enemmän kapasiteettia ja nopeutta kun on budjetoitu.
- **Luulet, että ongelmat ovat ohi kun tietovarasto on valmis ja käynnissä.** Tietovarasto- ja Business Intelligense projektit ovat jatkuvia prosesseja. Tietoa täytyy jatkuvasti lisätä tai muuttaa. Myös uusia analytiikka työkaluja täytyy mahdollisesti ottaa käyttöön projektin jälkeen.
- **Keskityt ad hoc tiedon louhintaan ja säännöllisiin raportteihin hälytysten sijaan.** Tiedon luonnollinen kulku tietovarastossa on:
  1. Tiedon poiminta operatiivisista järjestelmistä, tiedon puhdistaminen ja syöttäminen tietovarastoon
  2. Tukea ad hoc raportteja kunnes opit mitä ihmiset haluavat
  3. Muokkaat ad hoc raporteista säännöllisesti ajastettuja raportteja

Tämä ei kuitenkaan ole optimiratkaisu. Johtajat ja päälliköt ovat usein kiireisiä ja tarvitsevat aikaa lukeakseen raportteja. Hälytysjärjestelmät ovat parempi ratkaisu kuin ajastetut raportit. Hälytysjärjestelmä tarkkailee dataa ja lähettää tarvittaessa hälytyksen avainhenkilöille mikäli datassa tapahtuu ennalta määritelty kriittinen muutos.

Useissa yrityksissä tietovarastoprojekti onnistuu vain mikäli heillä on vahva keskijohdon tuki projektin toteutukselle. Tämä sama sääntö pätee myös muihin IT-projekteihin, mutta on erityisen tärkeä tietovarastoprojekteissa. (Turban ym. 2014:317-319)

Tiedon mallinnus ja käyttöoikeuksien rajaaminen on myös kriittinen tekijä projektin onnistumiselle. Tiedon mallinnuksessa on tärkeää, että juuri kyseisen osaston asiantuntijat ovat mukana määrittelemässä mitä tietoa kerätään ja missä muodossa se tallennetaan tietovarastoon. Tässä myös määritellään kenellä on oikeus päästä käsiksi dataan sen jälkeen kun se on tallennettu tietovarastoon.

## 2.6 Tietovaraston palvelun tarjoajat

Tietovarastoja on tarjolla useilta toimittajilta. Tietovarastoja vertailtaessa tulee ottaa huomioon palveluntarjoajan taloudellinen tilanne, linkitykset toiminnanohjausjärjestelmiin, konsulttien saatavuus, markkinaosuus, toimialakokemus ja kumppanuudet. Tietovarastotuotteita on kolmea tyyppiä, ensimmäisen tyyppisissä on toiminnot tiedon paikannukseen, lataamiseen, muokkaamiseen, puhdistamiseen sekä siirtämiseen ja lataamiseen tietovarastoon. Toinen tyyppi on sovellukset, jotka tallentavat ja hallinnoivat tietovaraston tietoa ja metatietoa. Kolmannen tyyppiset tuotteet tarjoavat loppukäyttäjälle työkalut tietovaraston tiedon analysointiin. Taulukossa 2 on listattu tietovarastotuotteiden palveluntarjoajat.

Taulukko 2. Tietovarastotuotteiden palveluntarjoajat. (Turban ym. 2014:310)

<b>Palvelun tarjoaja</b>	Tuotteen ominaisuudet
Computer Associates	Kattava kokoelma tietovarastotyökaluja ja tuotteita
DataMirror	Tietokannan hallinta ja suorituskyky työkaluja
Data Advantage Group	Metadata ohjelmisto
Dell	Tietokanta palvelimia
Embarcadero	Tietokannan hallinta ja suorituskyky työkaluja
Business Objects	Tiedon puhdistus (cleansing) työkalu
Harte-Hanks	Asiakkuuden hallinta tuotteita ja palveluja
HP	Tietokantapalvelimia
Hummingbird Ltd. (OpenText Connectivity Solutions)	Tietokannan hakukone ja tietovaraston tutkimus sovellus
Hyperion Solutions	Kattava kokoelma tietovarastotyökaluja ja tuotteita
IBM	Tietokanta työkaluja, tuotteita ja sovelluksia
Informatica	Tietokannan hallinta ja suorituskyky työkaluja
Microsoft	Tietokantatyökaluja ja tuotteita
Oracle	Tietokanta, ERP ja CRM työkaluja, tuotteita ja sovelluksia
SAS Institute	Tietokantatyökaluja, tuotteita ja sovelluksia
Siemens	Tietokanta palvelimia
SyBase	Kattava kokoelma tietovarastotyökaluja ja tuotteita
TeraData	Tietokantatyökaluja, tuotteita ja sovelluksia

### 3 TIEDON MUOKKAAMINEN

Operatiivisista järjestelmistä haetaan yleensä data ajastetusti yöaikaan, jolloin se ei rasi- ta yrityksen kriittisiä järjestelmiä. Datan hakua kutsutaan ETL-prosessiksi, Extract - Transform – Load eli poiminta, muokkaus ja lataus (Hovi ym. 2009: 23). Poiminta (Extract) vaiheessa data luetaan yhdestä tai useammasta tietokannasta. Poiminnan jälkeen data muokataan (Transform) haluttuun muotoon ja ladataan (Load) tietovarastoon. Nä- mä kolme vaihetta on integroitu yhteen ohjelmaan, joka hakee datan yhdestä tai use- ammasta tietokannasta ja siirtää ne tietovaraston tietokantaan. ETL-prosessi vie aikaa, joten reaaliaikaisessa tietovarastossa ei ole mahdollista tarkistaa, integroida, jalostaa eikä summata tietoja (Hovi ym. 2009: 30).

#### 3.1 ETL-prosessi

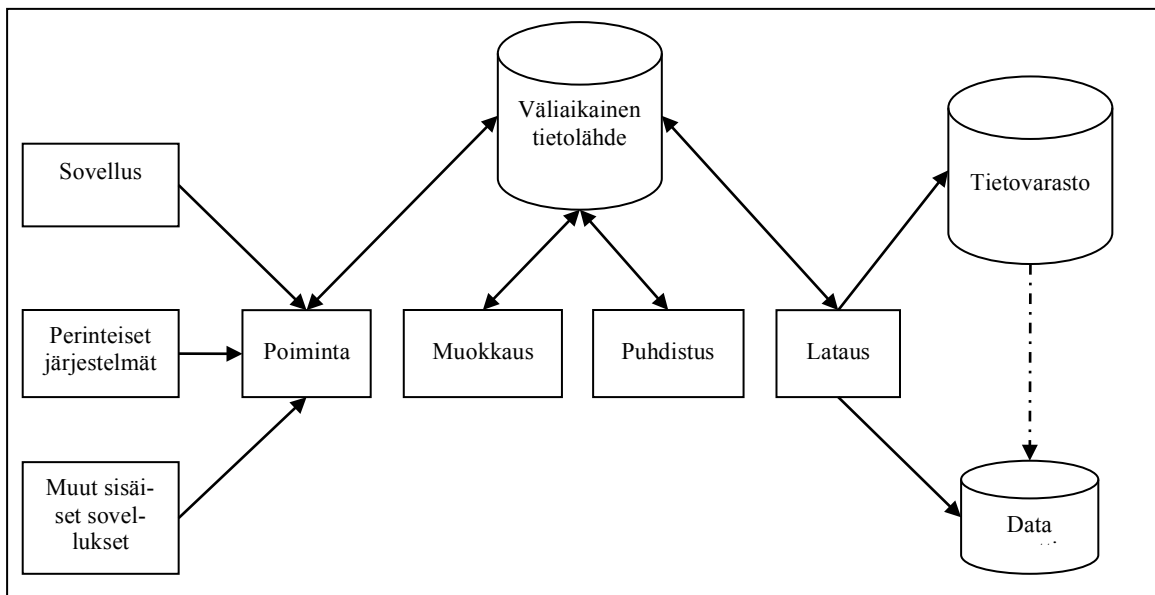
Datan prosessoinnissa tulee ottaa huomioon erityyppiset muuttujat. Eri tietokantoihin voidaan tallentaa tietoa eri tavalla, mutta tietovaraston tietokantaan vietäessä data pitää muuttaa samanmuotoiseksi. Esimerkiksi ”ei” voi olla ”0”, ”No” tai ”n” ja ”kyllä” voi olla ”1”, ”y” tai ”Yes”. Myös kategoriset muuttujat ovat ongelmallisia ja useimmiten niitä ei voi käyttää sellaisenaan tietovarasto ratkaisuisissa. Kategoriset muuttujat tulee muuttaa sarjaksi ”dummy” -muuttujia. Esimerkiksi muuttuja, joka voi saada arvot ”opiskelija”, ”työtön”, ”työssäkäyvä” tai ”eläkkeellä” muutetaan neljäksi eri muuttujak- si:

Opiskelija – Kyllä / Ei  
 Työtön – Kyllä / Ei  
 Työssä käyvä – Kyllä / Ei  
 Eläkkeellä – Kyllä / Ei

Näistäkin muuttujista voidaan jättää yksi muuttuja käyttämättä, koska mikäli muuttujis- ta tiedetään kolme, niin silloin tiedetään neljäskin muuttuja. Eli jos henkilö ei ole opis- kelija, työtön tai työssä käyvä niin silloin hänen on oltava eläkkeellä (Shmueli, Patel & Bruce 2010: s. 19).

ETL-prosessi (kuva 8) voidaan tehdä joko työntö- tai vetoperiaatteella. Työntöperiaat- teella tiedot poimitaan operatiivisesta järjestelmästä, muokataan oikeaan muotoon ja kirjoitetaan tiedostoihin. Tämän jälkeen tietoja muokataan vielä ennen kuin ne siirretään varsinaiseen tietovarastoon. Vetoperiaatteella toimiva ETL-prosessi lukee tiedot suoraan

operatiivisen järjestelmän tietokannasta. Tämä toteutus on yksinkertaisempi toteuttaa ja voi olla jopa joustavampikin kuin työntömenetelmä. Vetoperiaatteeeseen liittyy kuitenkin riskejä. Operatiivisen tietokannan päivitys voi olla kesken lukuhetkellä, jolloin tietovarastoon siirtyy väärää dataa tai jos operatiivinen järjestelmä vaihtuu, joudutaan implementoimaan koko ETL-prosessi uudelleen. (Hovi ym. 2009:61.)



Kuva 8. ETL-prosessi (Turban ym. 2011: 68).

### 3.2 ETL-työkalut

Datan prosessoimiseen operatiivisesta tietokannasta tietovarastoon voidaan tehdä erillisellä ETL-työkalulla tai tietokantatoimittajan toimittamalla työkalulla. Taulukossa 3 on listattu ETL-työkaluja ja niiden tarjoajia.

Yleensä tietovaraston tietokantatoimittaja toimittaa myös ETL-työkalun kustannus syistään. Erillisen työkalun ostaminen eri valmistajalta vaatisi ylimääräisten kustannusten lisäksi vielä IT-osaston konfiguroinnin. Kaupallisten ohjelmistojen lisäksi on tarjolla joitain Open Source –tuotteita, kuten Talend ja Pentaho. Useimmiten yritykset päätyvät vakiintuneisiin toimittajiin joilta voi ostaa koko tietovaraston ja jatkuvan tuen sen ylläpitoon. (Hovi ym. 2009:61.)

Taulukko 3. ETL-työkalut (Hovi ym. 2009: 60).

Tietovarasto	ETL-työkalu
Microsoft SQL Server	SQL Server Integration Services (SSIS)
Oracle	Oracle Warehouse Builder (OWB)
Oracle	Data Integrator
IBM DB2	Data Stage
IBM Cognos	Data Manager
SAP Business Objects	Data Integrator

ETL-työkalujen tarkoituksena on vähentää, tiedon siirtämiseen operatiivisesta järjestelmästä tietovarastoon, tarvittavaa ohjelmoinnin määrää. Työkalut tarjoavat graafisen käyttöliittymän sekä osaavat hakea tiedon erilaisista tietokannoista. Työkaluissa on myös ominaisuuksia tiedon tyyppimuunnoksille, yhdistelylle sekä tuplarivien poistoon. ETL-prosessi voidaan käynnistää joko ajastetusti tai herätepohjaisesti. Nykyään käytetään myös jatkuvaa latausta. Joihinkin ETL-tuotteisiin on lisätty profilointi ja monitorointi ominaisuuksia, jotka tarkkailevat tiedon laatuun liittyviä ominaisuuksia. ETL-työkaluissa on myös ominaisuus, jolla voidaan korjata virheitä tiedossa ja muokata raaka-dataa loppukäyttäjille valmiiseen muotoon. (Hovi ym. 2009:61, Xavier & Moreira 2013)

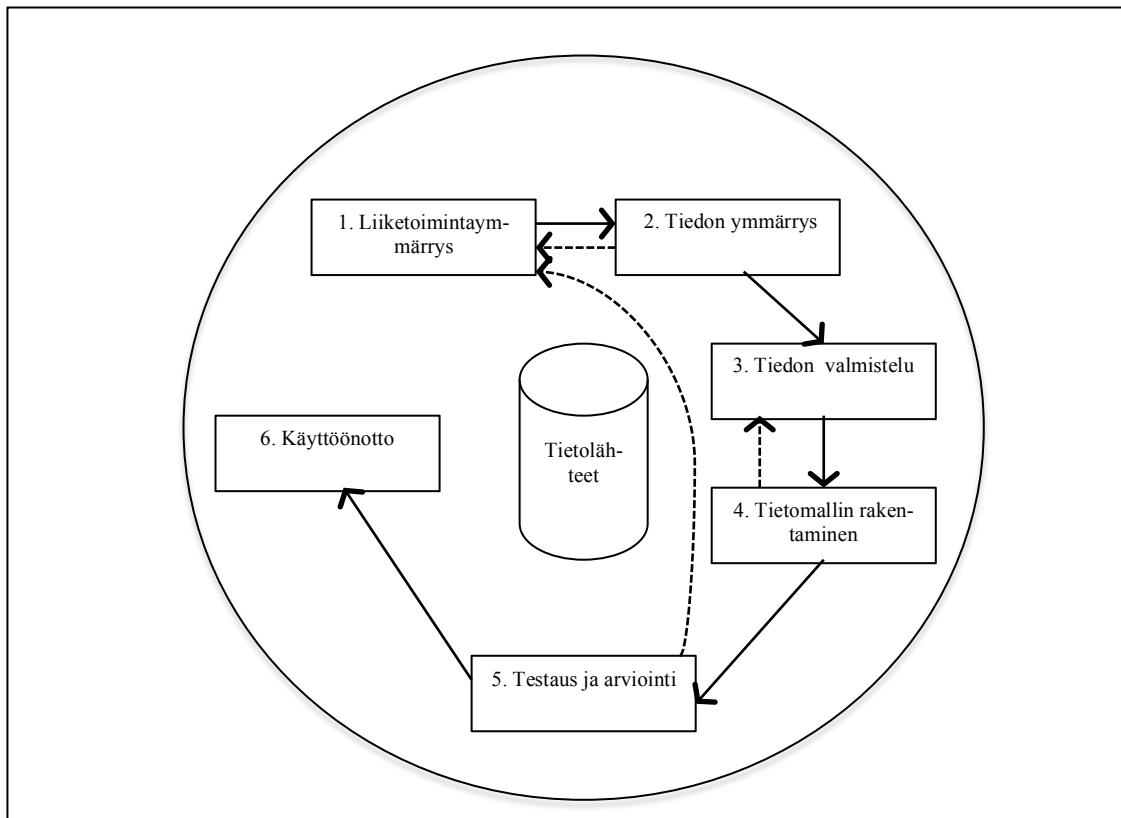
### 3.3 Tiedon louhinta

Tiedon louhinta (Data Mining) on prosessi, joka käyttää tilastollisia, matemaattisia ja tekoälyä tekniikoita kerätäksään tarvittavan tiedon suurista datamassoista. Hakuehdot, joilla tietoa kerätään voivat olla liiketoimintasääntöjä, yhtäläisyyksiä, korrelaatioita, trendejä tai ennusteita. Tiedon louhinta on määritelty seuraavasti: epätriviaali prosessi, jolla tunnistetaan validia, uutta, mahdollisesti hyödyllistä ja täysin ymmärrettäviä rakenteita (toistuvia merkkijonoja) tietokannassa. Tässä määrittelyssä avainsanat tarkoittavat:

- Prosessi - tiedon louhinta on sisältää useita iteraatioita
- Epätriviaali – joitain koe-tyyppisiä tai pääteltyjä hakuja joudutaan käyttämään
- Validi – löydetty tiedot ovat totuudenmukaisia riittäväällä varmuudella

- Uusi – löydetty tiedot eivät ole entuudestaan tunnettuja
- Mahdollisesti hyödyllinen – löydetystä tiedosta tulee olla hyötyä käyttäjälle
- Täysin ymmärrettävä – löydetyn tiedon tulee olla liiketoiminnallisesti järkevää

(Turban ym. 2014 s.196)



Kuva 9. Tiedon louhinnan kuusi vaihetta (Turban ym. 2014: 207).

Tiedon louhinta projekteissa käytetään yleensä standardoitua CRISP –prosessia (Cross-Industry Standard Process for Data Mining). Prosessissa on kuusi kohtaa, prosessi alkaa liiketoiminnan ymmärryksestä. Ennen kuin tiedon louhinta aloitetaan, tulee olla tiedossa mitä tietoa halutaan. Kohdassa kaksi tulee ymmärtää mitä tietoa ja mistä lähteistä tietoa halutaan kerätä.

Kolmantena tulee tiedon esivalmistelu, joka on CRISP -prosessin aikaa vievin vaihe. Tietojen mukaan tämä vaihe vie koko prosessin kestosta noin 80%. Tiedon esivalmistelun ensimmäisessä vaiheessa kerätään relevantti data tunnistetuista tietolähteistä. Esi-

valmistelun toisessa vaiheessa kerätty tieto puhdistetaan, tyhjät arvot täytetään sopivilla keskiarvoilla tai hylätään, ääriarvot poistetaan tai tasoitetaan muita arvoja vastaaviksi. Kolmannessa vaiheessa data muutetaan paremmin prosessoitavaan muotoon. Tiedon esivalmistelun viimeisessä vaiheessa datasta hylätään ylimääräinen osa. Vaikkakin dataa on hyvä olla paljon, myös liika data voi olla ongelma.

Prosessin neljännessä vaiheessa valitaan halutut mallinnus tekniikat liiketoiminnan tarpeita varten kolmesta pääkategorista: ennustaminen, yhdistäminen ja klusterointi. Ennustamisessa pyritään kertomaan tulevaisuudesta. Yhdistämisessä pyritään etsimään kiinnostavia yhteyksiä muuttujien välillä laajasta tietokannasta. Klusterointi taas jakaa asiakokonaisuudet segmentteihin niiden ominaisuuksien perustella. (Turban ym. 2014: 200)

Viidennessä vaiheessa edellisissä vaiheissa tehdyt mallit arvioidaan niiden tarkkuuden ja yleisluonteisuuden osalta. Kuudennessa eli viimeisessä vaiheessa on tiedon louhinnan käyttöönotto. Käyttöönotto voi yksinkertaisuudessaan olla pelkkä raportin luominen tai sitten monimutkainen toteutus koko organisaation laajuudessa toistettavasta tiedon louhinta prosessista. CRISP:n lisäksi muita tiedon louhinta prosesseja ovat SAS:n kehittämä SEMMA. (Turban ym. 2014: 207-212)

### 3.4 Big Data

Big Datalla tarkoitetaan tietoa, joka on peräisin lukemattomista tietolähteistä tietoverkosta. Big Data pitää sisällään miljardeja ja jopa biljoonia tietueita tietoa. Koska tämä tieto on usein huonosti strukturoitua ja epätäydellistä sellaisenaan, sen muokkaamiseen, analysointiin ja tiedon esittämiseen tarvitaan teknologiaa ja työkaluja. Big Datan ainoa haasteena ei ole tiedon tallentaminen ja säilyttäminen, vaan sen käsittely on osoittanut suuremmaksi haasteeksi. Tätä varten ollaan kehittämässä yhteistä käsitteellistä tietomallia nimeltä Linked Data (LDP). (Bansal & Kageman 2015)

Linked Data kuvaa tavan julkaista tietoa verkkoon, jolloin tieto on koneellisesti luettavissa. Tämä tarkoittaa, että tieto on yksiselitteisesti määritelty ja se voi olla linkitetty ulkoisiin tietolähteisiin. Ja kun ulkoinen tietolähde on tuotettu samaa tapaa noudattaen, näistä yhdessä muodostuu massiivinen tietovarasto. Linked Open Data (LOD) yhteisö on tuottanut suuren tietovaraston tämän tekniikan avulla. (Bansal ym. 2015)

Myös W3C on standardoinut arkkitehtuurin ja tavan lukea ja kirjoittaa Linked Data tietoja:

- Käytä asioiden nimenä URIa
- Käytä http URIa, jotta ihmiset voivat etsiä asioita käyttäen nimeä
- Kun joku etsii URIa, tarjoa hyödyllistä tietoa, käytä standardeja (RDF\*, SPARQL)
- Sisällytä tietoon linkkejä muihin osoitteita, jotta muut löytävät enemmän asioita

ETL-prosessilla muokataan tieto yhtenäiseen muotoon, tämä on normaalisti tehty oman yrityksen tarpeita varten. Big Datan tapauksessa tieto tulee muokata esimerkiksi Linked Data –muotoon ja laittaa muiden saataville (W3C 2015).

### 3.5 Metadata

Metadata on tietoa tiedosta. Metadata kuvailee tiedon rakennetta ja tarkoitusta. Harvat yritykset ymmärtävät metadatan merkityksen ja vielä harvemmat ymmärtävät kuinka luoda strategia metadatan suunnittelua ja toteutusta varten. Metadata määritellään käytön mukaan joko tekniseksi tai liiketoiminnalliseksi tiedoksi. Syntaktinen metadata sisältää kuvauksen datan syntaksista, eli missä muodossa data on tallennettu. Rakenteellinen metadata kuvailee datan rakennetta ja semanttinen metadata kertoo datan tarkoituksen. Metadata pitää sisällään tiedon, joka auttaa meitä ymmärtämään perinteistä dataa. (Turban ym. 2011: 56.)

Taulukossa 4 on esitetty erityyppisiä metadatoja ja mitä tietoa niihin tallennetaan. Staattinen metadata pysyy pidempään muuttumattomana, kun taas dynaaminen metadata kertoo järjestelmän tilasta. Tyypillisesti dynaaminen metadata kertoo esimerkiksi milloin järjestelmästä on otettu viimeksi varmuuskopio tai paljonko levyillä on tilaa jäljellä. Ilman metadataa loppukäyttäjien olisi vaikea löytää etsimäänsä dataa tietokannasta. Kun tiedetään mitä etsitään, niin etsittävä data löydetään helposti metadatan avulla. (Giovinazzo 2000: 44-45.)

Metadatalla on useita eri käyttäjiä. BI-ympäristössä loppukäyttäjä tarvitsee eniten metadataa löytääkseen haluamansa tiedon tietokannasta. Muita käyttäjiä ovat tietovarasto-arkkitehti, joka määrittelee tietovaraston rakenteen.

Taulukko 4. Metadata tyyppejä (Giovinazzo 2000: 44).

<b>Staattinen (rakenteellinen)</b>	
Nimi	Kertoo nimen millä dataelementti tunnetaan järjestelmässä. Esim. asiakas_nimi, asiakas_id
Kuvaus	Täydellinen kuvaus dataelementistä
Muoto	Datan esitysmuoto
Data tyyppi	Datan tyyppi, esim. kokonaisluku, boolean
Yhteydet	Määrittelee yhteyden järjestelmään, eli missä dataa käytetään
<b>Dynaaminen (operatiivinen)</b>	
Laatu	Kuvailee tiedon laadun järjestelmässä, kuten tarkkuus, täydellisyys, johdonmukaisuus
Tilastollinen	Kertoo mitä dataa käytetään eniten ja mitä osioita olisi syytä tehostaa järjestelmän suorituskyvyn parantamiseksi
Tila	Sisältää tiedon järjestelmän tilasta, sekä sen koska viimeksi on tehty varmuuskopiot

Tässä vaiheessa syntyy suurin osa metadatasta. Insinöörit parantavat tietokannan suorituskykyä tekemällä valmiiksi usein kysytyjä kyselyitä. Järjestelmä- ja tietokantavalvojat käyvät läpi tilastoja ja muokkaavat ympäristöä saavuttaakseen paremman suorituskyvyn. Kaikki tiedot näistä muutoksista tallennetaan metadata-tietoihin, joiden avulla loppukäyttäjät löytävät aina haluamansa tiedon. (Giovinazzo 2000: 46.)

### 3.6 Tietovarastoinnin apuvälineet

Perinteisellä arkkitehtuurilla toteutetun tietovaraston tieto pysyy muuttumattomana ETL-prosessin jälkeen. Master Data Managementin (MDM) avulla yritys voi hallita yrityksen ydintietoja. Ydintietoa on esimerkiksi asiakas-, tuote- ja organisaatitiedot. Master Data Managementin tavoite on koota useiden operatiivisten järjestelmien ydintieto yhteen paikkaan. Master Dataa hallitaan keskitetysti ja muut operatiiviset järjes-

telmät viittaavat Master Dataan, jolloin kaikissa järjestelmissä ydintieto on yhtenäistä. Master Data palvelee operatiivisten järjestelmien lisäksi myös tietovarastoa.

Enterprise Information Integration-työkaluilla (EII) haetaan tietoa operatiivisten järjestelmien tietokannoista ja yhdistellään sitä lennosta. EII-ratkaisut eivät yleensä korvaa tietovarastoa, koska niiden data on muokkaamatonta. EII-työkalujen avulla täydennetään tietovaraston tietoa reaaliaikaisella tiedolla. (Hovi ym. 2009:67)

## 4 BUSINESS INTELLIGENCE

Business Intelligence (BI) on termi, joka tarkoittaa eri asiaa eri ihmisille. Joillekin se tarkoittaa tiedon analysointia ja toisille arkkitehtuureja, työkaluja tai tietokantoja. Business Intelligencen päätavoite on mahdollistaa interaktiivinen yhteys tietovaraston tietoon ja mahdollistaa sen käsittely niin, että yrityksen päättäjät voivat tehdä tarpeellisia analyyskejä. Business Intelligence prosessin tarkoitus on muokata data informaatioksi, jonka avulla voidaan tehdä päätöksiä, jotka johtavat toimenpiteisiin. (Turban ym. 2011: 28-29.)

Business Intelligence keksittiin 1990-luvun puolivälissä, mutta konsepti on paljon vanhempi. Jo 1970-luvulla oli raportointijärjestelmiä, ne olivat staattisia ja kaksiulotteisia eikä niiden avulla ollut mahdollista analysoida tietoa. Myöhemmin järjestelmiin lisättiin ominaisuuksia, kuten dynaaminen moniulotteinen raportointi, ennustaminen, trendianalyysit ja porautuminen. Nämä ominaisuudet olivat vain ylimmän johdon käytettävissä. Myöhemmin nämä samat ominaisuudet ja paljon muita ominaisuuksia tulivat muidenkin käyttöön. Konseptin nimi Executive Information System muutettiin Business Intelligenceksi. Vuodesta 2005 lähtien Business Intelligence sovelluksiin on lisätty tekoälyä ja tehokkaita analysointia sisältäviä ominaisuuksia. (Turban ym. 2011: 28-29.)

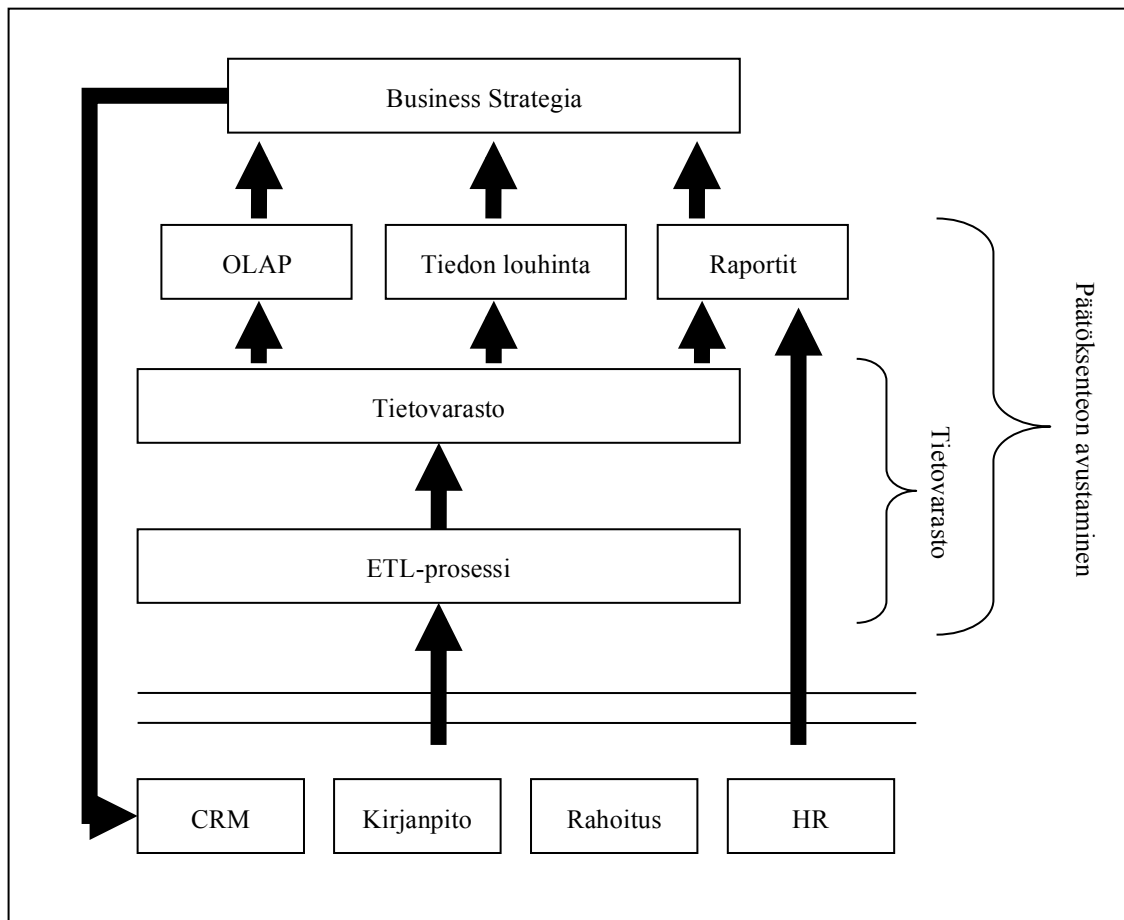
Business Intelligence kuvaa yrityksen toiminnasta kertyvää informaatiota, jonka avulla voidaan tehdä parempia päätöksiä ja ohjata toimintaa oikeaan suuntaan. Ilman BI-sovelluksia jokaisen tietoa tarvitsevan tulisi tuntea tietokantojen rakenne ja hakea tieto sieltä. BI-sovellukset esittävät tiedon visuaalisesti ja käyttäjät voivat hakea haluamansa tiedon ilman ohjelmointikokemusta. Kaikkien BI-sovellusten taustalla on tietovarasto, josta tieto haetaan. Reaaliaikainenkin BI-sovellus käyttää tietovarastoa, koska BI-sovelluksia varten data täytyy muokata oikeaan muotoon ennen esittämistä BI-sovelluksissa. (Hovi ym. 2009: 74.)

Business Intelligence tarjoaa yritykselle tarkkaa informaatiota, jota tarvitaan, kun tarkastellaan menestystä. Tämä tieto on tärkeää, jotta voidaan suunnitella tulevaisuutta ja tehdä tärkeitä päätöksiä. Yritykset on pakotettu tallentamaan, ymmärtämään ja hyödyntämään tietoa, jotta yritys voisi kehittää omia prosessejaan. Nykyään prosessit ovat nopeutuneet ja päätöksentekoaajat ovat lyhentyneet. Tämän vuoksi päätöksentekijät tarvitsevat oikean tiedon, oikealla hetkellä, oikeassa paikassa. (Turban ym. 2011: 32.)

Tietovarasto Instituutin (Data Warehouse institute), jossa on mukana 510 yritystä, mukaan Business Intelligencen käyttöönotto on säästänyt aikaa noin 61 %:lta yrityksistä.

Yrityksistä 57:a on pystynyt tekemään parempia strategisia päätöksiä ja 56%:a parempia taktisia päätöksiä. Lisäksi 39% yrityksistä on säästänyt rahaa Business Intelligencen käyttöönoton jälkeen. (Xavier ym. 2013)

Aikaisemmin Business Intelligencen hyödyntäminen toi yritykselle kilpailuedun. Nykyään BI-työkalujen puuttuminen on yritykselle haitaksi. BI sisältää liiketoiminnan kaikki osa-alueet, asiakashallinnan, kirjanpidon, rahoituksen, myynnin jne. Yrityksen toiminta aiheuttaa tapahtumia yrityksen operatiiviseen järjestelmään (kuva 10). Operatiivisesta järjestelmästä kerätään haluttu data, muokataan se oikeaan muotoon ja se tallennetaan tietovarastoon. Tietovaraston datasta voidaan tehdä erilaisia raportteja. Raporttien avulla yrityksen johto tekee strategisia päätöksiä, jotka vaikuttavat taas dataan, joka tallentuu operatiiviseen järjestelmään. BI prosessi on siis eräänlainen silmukka. (Giovinazzo 2000: 2.)



Kuva 10. Business Intelligence silmukka (Giovinazzo 2000: 2).

#### 4.1 Business Intelligence pilvessä

Viime vuosina pilvipalvelut ovat yleistyneet, uusia liiketoiminta muotoja ovat SaaS (software as a service), PaaS (platform as a service) ja IaaS (infrastructure as a service). Ensimmäinen tarkoittaa esimerkiksi verkossa toimivaa ohjelmistoa tai sovellusta, jonka käytöstä maksetaan. Tätä sovellusta ei siis ladata käyttäjän koneelle, vaan sovelluksen käyttö tapahtuu verkon kautta. PaaS on palvelu, jossa esimerkiksi yritykselle tarjotaan alusta tai ympäristö, jossa kehittää, suorittaa ja hallinnoida web-sovelluksia. IaaS taas tarjoaa käyttäjälle koko infrastruktuurin käytettäväksi. Tämä sisältää verkkoyhteydet, tallennustilan, palvelimet ja niiden ylläpidon. (Al-Aqrabi, Liu, Hill & Antonopolous 2014)

Vaikka BI järjestelmä olisi tällä hetkellä toimiva, liiketoiminnan omistajien täytyy etsiä vaihtoehtoja hallitsemattoman järjestelmä- ja tallennus resurssien kasvun vuoksi itse ylläpidettyihin järjestelmiin. Jossain vaiheessa järjestelmien ylläpito ja päivitys tulee kohtuuttoman kalliiksi. Tällöin pilvipalvelut ovat hyvä ratkaisu BI järjestelmälle, jolloin järjestelmien ja infrastruktuurin ylläpito on palveluntarjoajalla. Pilvipalvelut ovat myös hyvä ratkaisu pienemmille yrityksille, jolloin heidän ei tarvitse rakentaa ja ylläpitää itse BI järjestelmien vaativaa infrastruktuuria. (Al-Aqrabi ym. 2014)

Yksinkertainen BI-järjestelmä koostuu loppukäyttäjistä, raportointipalvelimesta, tietovarastosta, ETL-palvelimesta ja operatiivisista järjestelmistä. Tällaisessa BI-järjestelmässä laitteisto ja ohjelmisto kulut voivat nousta jopa miljooniin euroihin. Pilvessä olevassa BI-järjestelmässä vain loppukäyttäjä ja operatiivisten järjestelmien laitteistot ja ohjelmistot ovat yrityksen vastuulla. (Gash, Ariyachandra & Frolick 2011, )

Pilvessä olevaa Business Intelligence kutsutaan BIaaS (Business Intelligence as a service). Pilveen rakennettu BI-järjestelmä tuo yritykselle kustannussäästöjä laitteistojen ja ohjelmistojen lisäksi myös järjestelmän suunnittelu- ja toteutusvaiheessa. Yleensä järjestelmän toimittavalla yrityksellä on jo valmis perussuunnitelma järjestelmästä ja ympäristöstä, ja järjestelmä sitten räätälöidään yritykselle sopivaksi. (Gash ym. 2011, Chang 2014)

## 4.2 Business Intelligence sovellukset

Business Intelligence-sovelluksia on saatavilla useilta eri valmistajilta. Isoimmat valmistajat ovat Teradata, MicroStrategy, Microsoft, IBM, Oracle, SAS ja SAP. Useimmat edellisistä ovat kasvaneet suuriksi toimijoiksi yritysostojen kautta. Koska BI-sovellus tarvitsee yleensä tietovaraston toimiakseen, usein ostetaan sekä tietovarasto että BI-sovellus samalta toimittajalta. Tämä on oikeastaan ainoa käytetty vaihtoehto kustannussyistä. BI-sovellukseen kuuluu useita muitakin tuotteita tietovaraston ja itse BI-sovelluksen lisäksi. Usein sovelluksen mukana toimitetaan mm. raportointityökaluja, päätöksentekoaupuvälineitä, portaaleja, analysointityökaluja, tuloskortteja ja mittaristoja. (Turban ym. 2011:41.)

Taulukossa 5 on listattu tunnettuja BI-työkalujen valmistajia sekä mainittu työkalujen nimet. Suurten BI-työkalujen valmistajien lisäksi markkinoilla on myös pienempiä haastajia.

Taulukko 5. Business Intelligence-työkalut (Turban ym. 2011:41).

<b>Valmistaja</b>	<b>Työkalu</b>
IBM	IBM Cognos Business Intelligence
Microsoft	Power BI
Microstrategy	Microstrategy
Oracle	Business Intelligence Enterprise Edition
Oracle	Oracle Hyperion System
Pentaho BI Suite	Pentaho
QlikTech	QlikView
SAP	SAP BusinessObjects
SAP	SAP NetWeaver BI
SAS	SAS Enterprise BI Server
Tableau Software	Tableau
WebFOCUS	Information Builders

BI-sovelluksen mukana toimitetaan yleensä myös tietovarasto, mutta useimmat työkalut on mahdollista integroida myös muiden valmistajien tietokantoihin. BI-sovellukset in-

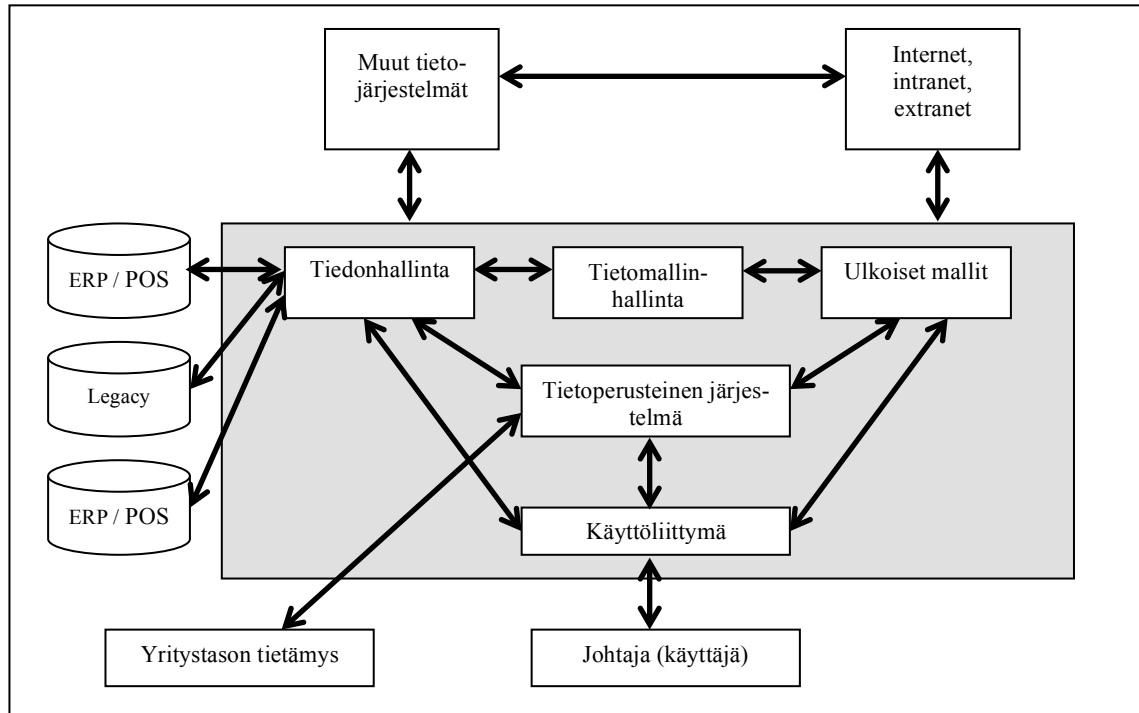
tegroidaan usein myös toiminnanohjausjärjestelmään, jolloin saadaan reaaliaikaista tietoa käyttöön. Yleisimmin ERP/BI integraatiota on käytetty taloudellisten päätösten avuksi, mutta nykyään sitä käytetään myös markkinoinnin, henkilöstöhallinnon ja tuotannon päätöksenteon apuvälineinä. (Turban 2011: 261.)

### 4.3 Päätöksenteon apuvälineet

Päätöksenteon apuvälineet (Decision Support System, DSS) on yleensä tehty jotain tiettyä ongelmaa varten tai arvioimaan tiettyä mahdollisuutta. Raportointi näyttää BI-järjestelmissä suurta roolia. BI-järjestelmät tarkkailevat tilanteita ja tunnistavat ongelmia ja mahdollisuuksia käyttäen analyttisiä menetelmiä. BI-järjestelmiin on sisällytetty tietomallit ja tietokantayhteys. BI keskittyy siis raportointiin ja etsimään ongelmia. Päätöksenteon apuvälineet käyttävät taas omaa tietokantaa ja ne on kehitetty ratkaisemaan BI:n löytämiä ongelmia. Päätöksenteon apuvälineet kehitetään varta vasten löydettyjä ongelmia varten. Näitä apuvälineitä voi käyttää yksi käyttäjä tietokoneella tai se voi olla internetin kautta käytettävä sovellus, jolla voi olla useita yhtäaikaista käyttäjiä. (Turban, Sharda & Delen 2011: 75.)

Web-pohjaisessa DSS-arkkitehtuurissa prosessointi on hajautettu usealle palvelimelle. Sovelluspalvelimella sijaitseva sovellus käynnistetään web-selaimen kautta. Sovellus muodostaa yhden tai useamman mallin omasta tietokannasta ja tietomalleistaan. Sovellus voi myös hakea tietoa tietovarastosta tai operatiivisesta järjestelmästä. Mikäli käyttäjä vaatii mallin optimointia, tietomalli tietoineen siirretään optimointipalvelimelle. Optimointi palvelin hakee tietovarastosta tarvittaessa lisää tietoa. Tämän jälkeen optimointipalvelin ratkaisee ongelman ja lähettää vastauksen suoraan web-selaimelle. Sovelluspalvelin muokkaa muodostetut ratkaisuraportit johtajille luettavaan muotoon ja lähettää ne automaattisesti sähköpostilla tai julkaisee ne portaalissa. (Turban ym. 2011: 75-76.)

Päätöksenteon apuväline voi olla monimutkainen ja useasta eri komponentista koostuva järjestelmä. Määrittelyjen mukaan päätöksenteon apuväline tarvitsee vähintään kolme komponenttia: tietokannanhallintajärjestelmän, tietomallinhallintajärjestelmän sekä käyttöliittymän. Kuvassa 11 on kuvattu päätöksenteon apuväline kaikkine komponentteineen. Useimmat näistä web-järjestelmistä on toteutettu Javalla ja ne pyörivät sovelluspalvelimilla.



Kuva 11. Päätöksenteon järjestelmä (Turban ym. 2014: 85).

Järjestelmään voi kuulua:

**Tiedonhallintajärjestelmä (The Data Management Subsystem)** sisältää tietokannan, jota ylläpidetään tietokannanhallintaohjelmiston (DBMS) avulla ja on monesti kytketty yrityksen tietovarastoon.

**Tietomallinhallintajärjestelmä (The Model Management SubSystem)** on sovellus, joka pitää sisällään taloudelliset, tilastolliset, hallinnolliset tai muut kvantitatiiviset mallit, jotka mahdollistavat tietojen analysoinnin. Tämä komponentti on yleensä kytköksissä yrityksen sisäisiin tai ulkoisiin tietomalleihin. Tätä kutsutaan usein myös tietoperusteinen hallintajärjestelmä (Model Base Management System, MBMS).

**Käyttöliittymä (The User Interface Subsystem)** on rajapinta, jonka avulla päätöksenteonapuvälinettä käytetään. Usein käyttöliittymä on web-sivu, jolloin käyttäjälle tarjotaan graafinen käyttöliittymä. Joskus käyttöliittymä voi olla myös excel-asiakirja.

**Tietoperusteinen hallintajärjestelmä (The Knowledge-Based Management Subsystem)** voi tukea kaikkia muita järjestelmän komponentteja tai se voi toimia itsenäisenä järjestelmänä (Turban ym. 2014: 85-86).

Päätöksenteon apuvälineitä ei ole määritelty yksimielisesti, eikä sen ominaisuuksia ja kykyjä ole määritelty tarkasti. Seuraavassa listassa on kuitenkin joitain DSS:n ominaisuuksia:

1. Osittain jäsennellyt tai jäsentymättömät ongelmat
  2. Palvelee johtajia kaikilla tasoilla
  3. Palvelee yksilöitä ja ryhmiä
  4. Toisistaan riippumattomat tai peräkkäiset päätökset
  5. Tukee älykästä suunnittelua, valintoja ja toteutusta
  6. Tukee erilaisia päätöksenteko prosesseja ja tyylejä
  7. Mukautuva ja joustava
  8. Interaktiivinen ja helppokäyttöinen
  9. Tehokas
  10. Ihmiset ohjaavat prosessia
  11. Loppukäyttäjien helppo kehittää
  12. Mallinnus ja analyysit
  13. Tietokantayhteys
  14. Erillisjärjestelmä, integroitu tai web-pohjainen
- (Turban ym. 2011: 77.)

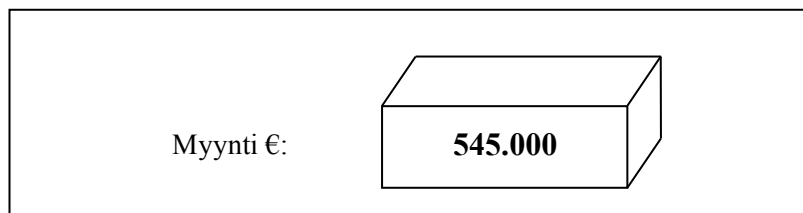
Automaattisia päätöksentekojärjestelmiä käytetään usein tuotteiden tai palveluiden hinnoittelussa. Lentolippujen hinnat on hyvä esimerkki: lentoyhtiön myy lippuja ja mikäli vain 70 % lennon paikoista on myyty tiettyä ajankohtana ennen lentoa, annetaan x prosentin alennus muille kuin business-luokan matkustajille. Tällaiset päätökset on tehty aikaisemman datan analysoinnin perusteella ja syötetty operatiiviseen järjestelmään matemaattisten mallien avulla. (Turban ym. 2011: 28-29, 34.)

Tapahtumapohjaiset hälytykset varoittavat BI-käyttäjää epätavallisista tapahtumista. Esimerkiksi luottokorttiyhtiö saa hälytyksen, mikäli luottokortilla yritetään maksaa epätavallisen suuri summa tai yritetään maksaa epätavallisessa sijainnissa. (Turban ym. 2011:35.)

#### 4.4 Raportointi

Business Intelligence -järjestelmän tarkoituksena on tuottaa raportteja käyttäjilleen. Raporttien informaation tulee olla oikeaa ja nopeasti saatavilla. Raportit luodaan tietovarastotietokannasta tai datamartista, lukuun ottamatta reaaliaikaista BI-ratkaisua, jossa data haetaan operatiivisesta järjestelmästä. Uusi raportti luodaan tekemällä kysely SQL-kielellä tietolähteestä. Usein kyselyn tekemistä varten on luotu käyttöliittymä, jossa käyttäjä valitsee halutut tiedot raporttiin ja sovellus muodostaa SQL-lausekkeen käyttäjän puolesta. Vaikka uusien raporttien luominen on tehty helpoksi, usein tietohallinto tai ulkopuolinen konsultti muodostaa uudet raportit. Raportteihin voidaan sisällyttää muuttuvia elementtejä, jolloin useat eri käyttäjät voivat hyödyntää samaa raporttia. Raporttiin syötetään hakuehdot ennen varsinaista tiedon hakua. Raporttien hakuehdot voidaan myös automatisoida, jolloin hakuehdot määräytyvät esimerkiksi sisään kirjautumisen yhteydessä.

Moniulotteinen analysointi (OLAP, On-line Analytical Processing) on menetelmä, jolla on helppoa ja tehokasta toteuttaa liiketaloudellisia analyyseja. Kuvassa 12 on OLAP-kuutio ilman ulottuvuuksia. Kuutio voi esittää esimerkiksi yrityksen kokovuoden myyntiä kaikkien tuotteiden osalta.



Kuva 12. OLAP-kuutio. Nolla ulottuvuutta (Hovi 2009:92).

Kuvassa 13 kuutioon on lisätty ensimmäinen ulottuvuus, kvartaali. Nyt kuutiosta saadaan selville koko vuoden myynti kvartaaleittain sekä aggregaattina koko vuoden myynti.

Yht.	150.000	125.000	130.000	140.000	<b>545.000</b>
	Q1	Q2	Q3	Q4	Vuosi 2007
X	60.000	35.000	40.000	30.000	<b>165.000</b>
Y	50.000	65.000	20.000	40.000	<b>175.000</b>
Z	40.000	25.000	70.000	70.000	<b>205.000</b>
Yht.	<b>150.000</b>	<b>125.000</b>	<b>130.000</b>	<b>140.000</b>	<b>545.000</b>
	Q1	Q2	Q3	Q4	Vuosi 2007

Kuva 13. Ylempänä yksiulotteinen ja alempana kaksiulotteinen OLAP-taulu (Hovi ym. 2009: 91).

Toinen ulottuvuus voi olla esimerkiksi tuote. Nyt kuutio näyttää taulukolta (kuva 13), jossa jokainen rivi on oma tuotteensa ja jokainen sarake edustaa kvartaalia. Viimeisestä sarakkeesta jokaiselta riviltä löytyy tuotteen kokonaismyynti ja alimmalta riviltä jokaisesta sarakkeesta löytyy kvartaalin kokonaismyynti yhteenlaskettuna. Kuvassa 14 on kuvattu kolmiulotteinen OLAP-kuutio. Kolmantena ulottuvuutena kuutioon on lisätty myyntialue. Nyt OLAP-kuutio pitää sisällään tiedot yrityksen myynneistä kvartaaleittain, tuotteittain sekä alueittain. OLAP-kuutioiden ulottuvuuksia voi olla rajattomasti, mutta yli 10 ulottuvuuden kuutiot on todettu vaikeaksi hallita. Kuutioon ulottuvuuksien tarkasteleminen vaatii porautumista. Esimerkiksi jos halutaan tarkastella tietyn kvartaalin osalta tietyn tuotteen myyntiä, tulee ensin porautua kvartaalin ja tuotteen kohdalta, jolloin näkyviin tulee tuotteen maantieteelliset myynnit.

	<i>Itä</i>	35.000	10.000	15.000	20.000	80.000
	<i>Länsi</i>	15.000	5.000	15.000	5.000	40.000
	<i>Pohj.</i>	10.000	20.000	10.000	5.000	45.000
X		60.000	35.000	40.000	30.000	<b>165.000</b>
Y		50.000	65.000	20.000	40.000	<b>175.000</b>
Z		40.000	25.000	70.000	70.000	<b>205.000</b>
Yh		<b>150.000</b>	<b>125.000</b>	<b>130.000</b>	<b>140.000</b>	<b>545.000</b>
		Q1	Q2	Q3	Q4	Vuosi 2007

Kuva 14. Kolmiulotteinen OLAP-kuutio (Hovi 2009:92).

#### 4.5 Mittaristot

Business Intelligencen tuottamaa dataa seurataan erilaisilla mittareilla. Liiketoiminnan avainindikaattoreiden (KPI, Key Performance Indicators) avulla seurataan yrityksen suoriutumista johdon asettamista strategisista tavoitteista. Esimerkkinä strateginen tavoite voi olla markkinaosuuden kasvattaminen, jolloin KPI voisi olla myynnin kasvattaminen. Myyntiluku voidaan jakaa vielä pienempiin osiin, jolloin voidaan seurata tietyn tuoteryhmän myynnin kehitystä tai vaikka yhden tuotteen myyntiä. BI:n käyttäjät valitsevat seurattavat mittarit, joita voi helposti seurata kojelaudasta (Dashboard). Kojelaudasta näkee yhdellä silmäyksellä kaikkien seurattavien mittareiden tilanteen.

Muita mittareita ovat Business Activity Monitoring (BAM), joka toimii kuten Dashboard mutta pyrkii reaaliaikaisuuteen. BAM on tarkoitettu nykytilan ymmärtämiseen toisin kuin Dashboard, jossa dataa verrataan aikaisempaan dataan. Kaplan ja Nortonin Balanced Scorecardia (BSC) käytetään myös mittarina. BSC jakaa seurantakohteet neljään osaan: talous, asiakkaat, sisäiset prosessit ja oppiminen. Näihin osa-alueisiin kohdistetaan mitattavia tunnuslukuja, joita seurataan. Aiemmin BSC-ratkaisuihin kerättiin data manuaalisesti neljännesvuosittain, mutta nykyään aikaväli on liian pitkä. BSC:tä on yritetty tuoda osaksi BI:tä, mutta tulokorttien kytkennästä neljään osa-alueeseen on luovuttu ja siirrytty seuraamaan yritykselle tärkeitä tunnuslukuja. (Hovi ym. 2009:100.)

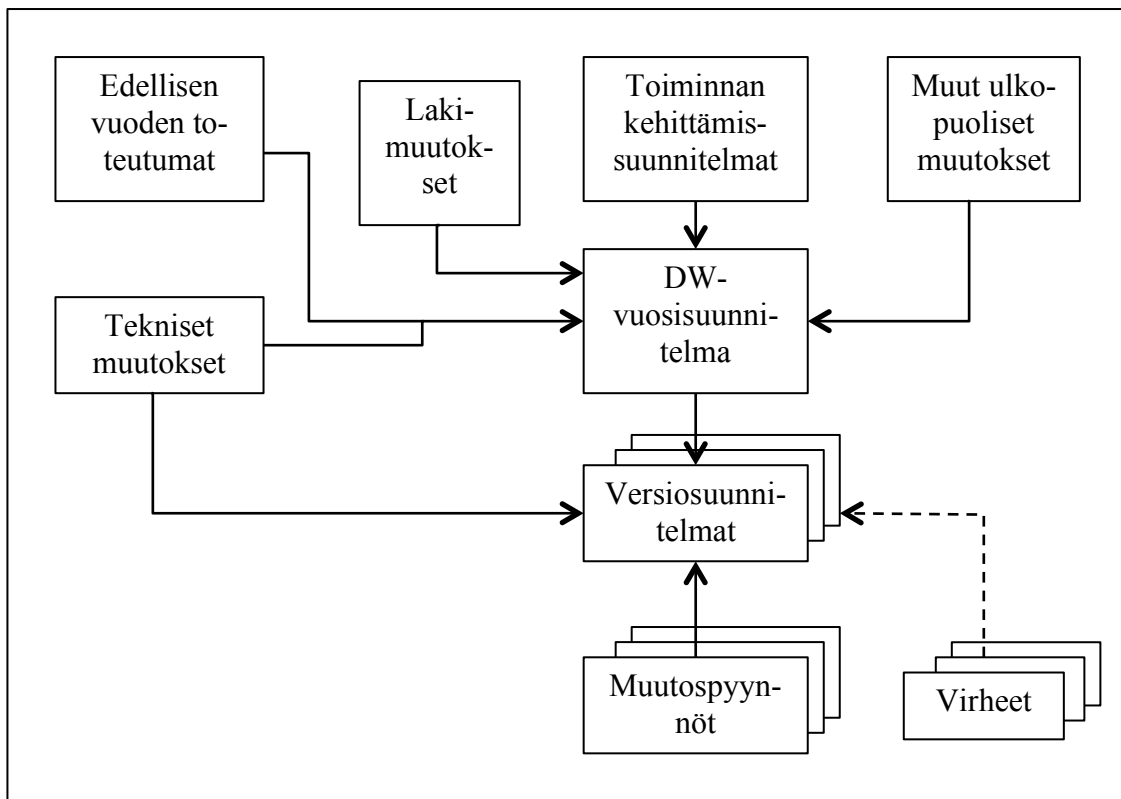
#### 4.6 BI-projektin toteutus

BI-projektin toteutus on hyvin monimutkainen prosessi, jonka onnistumista voidaan mitata monella eri tavalla. Saavatko oikeat ihmiset tietoa oikeaan aikaan? Voidaanko informaatiota hyödyntää paremmin? Jotta näihin kysymyksiin voitaisiin vastata, vaatisi se BI-projektin onnistuneen läpiviennin. BI-projektin onnistumiseen vaikuttaa monta eri tekijää, joista kriittisimmät ovat:

- Liiketoimintalähtöinen metodologia ja projektin johtaminen
- Selkeä visio ja suunnittelu
- Sitoutunut johdon tuki ja rajoitus
- Tiedon hallinta ja laatutekijät
- Ratkaisujen suunnittelu käyttäjien vaatimusten mukaisiksi
- Suorituskykytekijät
- Kestävä ja laajennettava ympäristö

Edellä lueteltujen asioiden huomioonottaminen BI-projektin suunnitteluvaiheessa luo hyvät valmiudet projektin onnistumiselle.

BI-projekti tulisi toteuttaa vain ja ainoastaan liiketoiminnan tarpeisiin. Kaikki raportit ja mittarit tulisi perustua loppukäyttäjän tarpeisiin ja IT-osaston tulisi olla projektissa vain toteuttajana ja teknisenä tukena. Loppukäyttäjien tulisi olla projektissa mukana jo suunnitteluvaiheessa. Projektia ei tulisi myöskään aloittaa, ellei yrityksellä ole selkeää visiota tai suunnitelmaa projektin toteuttamiseksi. BI on liian suuri ja monimutkainen projekti toteutettavaksi ilman todellista tarvetta. Johdon tuki on BI projektille välttämätön, jotta projekti saa riittävät resurssit käyttöönsä. BI on pitkäkestoinen prosessi, jota kehitetään koko ajan tuottamaan uutta informaatiota. BI ei ole siis projekti, joka alkaa ja loppuu, vaan se vaatii jatkuvaa huomiota. (Hovi ym. 2009: 121-124, Turban ym. 2011: 255-256.)



Kuva 15. Tietovaraston kehityssuunnitelma. (Hovi ym. 2009: 180)

Tietovaraston jatkokehittämisen tulee olla suunniteltua vuositasolla sekä versiotasolla. Vuosisuunnitelmassa määritellään miten tietovarastoa kehitetään seuraavan vuoden aikana. Kuvassa 15 on kuvattu tietovaraston kehityssuunnitelmassa huomioonotettavia asioita. Tämä suunnitelma tehdään yhdessä liiketoiminnan vuosisuunnitelman ja IT-osaston muun vuosisuunnitelman kanssa, jotta tietovaraston kehitys kulkee linjassa muiden kehitystoimien kanssa. (Hovi ym. 2009: 180)

Vuosisuunnitelman laatimisessa käytetään apuna edellisen vuoden toteumia ja liiketoiminnan suunnitelmia. Myös viranomaisten vaatimukset tulee ottaa huomioon. Näitä ovat esimerkiksi lainsäädännön muuttuminen ja EU vaatimukset. Asiakkaat voivat myös vaatia tarkempaa informaatiota ostokäyttäytymisestään, jolloin tietovarasto on paras paikka tällaisen tiedon säilytykseen. Vuosisuunnitelmaa tehtäessä tulee myös huomoida yrityksen muut tekniset suunnitelmat, jotta järjestelmät ovat jatkossakin yhteensopivia keskenään. (Hovi ym. 2009: 181)

Vuosisuunnitelman lisäksi laaditaan versiosuunnitelma. Versiosuunnitelmassa on kuvattu yhdessä versiossa tuotantoon vietävä kokonaisuus. Versioiden lukumäärä tavallisesti on 2-5 versiota vuodessa. Tämä luku on tietenkin tapauskohtainen ja monesti versioita joudutaan lisäämään esimerkiksi kriittisen tietoturvapäivityksen tai virheen vuoksi. Versiosuunnitelmaa laatiessa käytetään apuna samoja asioita kuin vuosisuunnitelmaa laatiessa. Näiden lisäksi suunnitelmaan lisätään yksittäiset muutospyynnöt sekä virheet. Yksittäiset muutospyynnöt ovat liiketoiminnalta käytön myötä tulleita muutospyyntöjä. Virheitä taas esiintyy järjestelmissä aina. Jotkut virheet on korjattava heti, ja jotkut taas ovat sellaisia, että ne eivät haittaa käyttöä ja voidaan korjata myöhemmin. Nämä virheenkorjaukset viedään versiossa tuotantoon. (Hovi ym. 2009: 181-182)

Tietovarasto vaatii jatkuvasti muutoksia ja uusia ominaisuuksia. Uudet ominaisuudet voidaan tehdä joko pienkehityksenä tai versiossa. Yleensä suuremmat muutokset kannattaa viedä versioissa. Pienempiä ja toisista järjestelmistä riippumattomia muutoksia voidaan viedä pienkehityksenä. Kehitystöitä ei voida yksiselitteisesti määrittellä versio- tai pienkehitystöiksi. Tietovaraston, kuten muidenkin järjestelmien jatkokehitystä ja ylläpitoa on organisoitava. Organisointimalleja on useita, mutta tässä esitellään yksi vaihtoehto. Organisaatioon kuuluu:

- **Tietovaraston koordinoitiryhmä** vastaa tietovaraston kokonaisuudesta ja sen toimivuudesta.
- **Tietovarastoarkkitehti** on vastuussa järjestelmien yhteensopivuudesta ja suunnittelusta.
- **Tietovastaavat** ovat vastuussa tietovarastossa olevan tiedon määrittelystä ja kuvauksesta. Tietovastaavat toimivat myös käyttäjien apuna.
- **Tietovaraston tekninen tuki** vastaa tietovaraston teknisestä toimivuudesta. Nämä henkilöt ovat käyttäjien tukena sovelluksiin liittyvissä ongelmissa. Tietoon liittyvissä ongelmissa auttaa tietovastaavat.
- **Osaamiskeskuksissa** on sekä tietovarastoinnin että liiketoiminnan asiantuntijoita. Tietovaraston jatkokehitys on tämän yksikön vastuulla.

(Hovi ym. 2009: 183)

## 5 LIIKETOIMINNAN SUORITUSKYVYN JOHTAMINEN

Liiketoiminnan suorituskyvyn johtaminen (Business Performance Management, BPM) viittaa liiketoiminnan prosesseihin, metodologioihin, mittareihin ja teknologioihin joilla mitataan, valvotaan ja johdetaan liiketoiminnan suorituskykyä. Liiketoiminnan suorituskyvyn johtamisesta on käytetty myös termiä yrityksen suorituskyvyn johtaminen (Corporate Performance Management, CPM tai Enterprise Performance Management, EPM) tai strateginen yrityksen johtaminen (Strategic Enterprise management). (Turban ym. 2014:341)

### 5.1 Liiketoiminnan suorituskyvyn johtamisen eroaminen BI:stä

Kun liiketoiminnan suorituskyvyn johtaminen esiteltiin ensimmäisen kerran erillisenä konseptina, sen ero BI:een oli epäselvä. Aluksi luultiin, että BPM on vain uusi termi BI:lle tai että BPM on BI:n seuraava sukupolvi. Edelleen näiden kahden konseptin ero on epäselvä, koska BI:tä ja BPM:a myydään samojen palvelujen tarjoajien toimesta ja BI on ratkaiseva tekijä BPM:n toteuttamisessa. (Turban ym. 2014:341)

Business Intelligence kuvailee käytetyt teknologiat, joita tarvitaan yrityksen kannalta tärkeän tiedon analysointiin ja raportointiin. BPM tarkemmin määriteltynä on BI + suunnittelu. BPM sisältää lähes aina BI:n käytännöt ja ohjelmistot. BPM ei kuitenkaan rajoitu tähän, vaan sisältää organisaation laajuisen strategian estääkseen paikallisuten yksiköiden optimoinnin koko organisaation suorituskyvyn kustannuksella. BPM on jatkuva kokoelma prosesseja, onnistuessaan sillä on vaikutus koko organisaatioon. (Turban ym. 2014: 342)

BPM prosessi alkaa strategian muodostuksella ja tavoitteiden asettamisella. Seuraavaksi suunnitellaan budjetti ja ennusteet sekä luodaan suunnitelmat tavoitteiden saavuttamiseksi. Tämän jälkeen seurataan tilannetta ja analysoidaan toteumia ja tarvittaessa tehdään korjaavia toimenpiteitä.

1. **Muodosta strategia.** Missio, arvot, tavoitteet, kannustimet, strategia kartat.
2. **Suunnittele.** Budjetti, suunnitelmat, ennusteet, mallit, aloitteet, määränpää.
3. **Seuraa ja analysoi.** Suorituskyky koetaulut, raportit, analyttiset työkalut.
4. **Toimi ja mukauta.** Tulkitse, tee yhteistyötä, arvioi, päätä, toimi, mukauta, seuraa (Turban ym. 2014:343)

## 5.2 Strategia

Termillä strategia on monta merkitystä, mutta useimmiten se tarkoittaa ”Missä haluamme olla tulevaisuudessa?”. Useimmissa yrityksissä vastaus tähän kysymykseen on esitetty strategiasuunnitelmassa. Tyypillisesti strategiakartta alkaa ylhäältä ja siinä on kuvattu ensimmäisenä organisaation laajuinen näkymä. Yleisiä strategiasuunnittelun vaiheista ovat:

### 1. Suorita nykytila-analyysi

- Tämä analyysi selvittää yrityksen nykytilan ja vastaa kysymykseen ”Missä me olemme nyt?”

### 2. Määrittele ajanjakso, jolle strategia suunnitellaan

- Perinteisesti yritykset tekevät suunnitelmat vuosittain kolmeksi – viideksi vuodeksi eteenpäin. Suunnitelmien pituus vaihtelee markkinoiden vaihtelun ja ennustettavuuden mukaan. Mitä suurempi vaihtelu, sitä pienempi ennustettavuus ja sitä lyhyemmäksi ajaksi suunnitelmat voidaan tehdä.

### 3. Analysoi toimintaympäristö (esim. SWOT analyysi)

- SWOT-analyysissä tunnistetaan mahdolliset ja todelliset yritykseen vaikuttavat tekijät, kuten: avainasiakkaat, markkinat, kilpailijat, hallitus ja toimiala

### 4. Määrittele kriittiset menestys tekijät

- Määritellään ne tekijät, joissa yrityksen täytyy menestyä, jotta voidaan menestyä markkinoilla. Näitä ovat esimerkiksi tuotteita valmistavassa yrityksessä tuotteiden laatu ja tuote innovaatiot

### 5. Suorita puuteanalyysi

- Tämä analyysi tunnistaa ja priorisoi sisäiset vahvuudet ja heikkoudet yrityksen prosesseissa, rakenteissa, teknologioissa ja sovelluksissa.

### 6. Luo strateginen visio

- Strateginen visio tarjoaa kuvan tai mielikuvan siitä, miltä yrityksen tulisi näyttää tulevaisuudessa.

## 7. Luo liiketoiminta strategia

- Liiketoiminta strategia tulee luoda edellisistä kohdista saatujen tietojen perusteella. Strategian tulisi olla linjassa yrityksen vahvuuksien kanssa, käyttää hyödyksi yrityksen mahdollisuuksia, osoittaa heikkoudet ja vastata uhkiin.

## 8. Tunnista strategiset tavoitteet

- Tässä kohdassa suuntaa antavat strategiset tavoitteet määritellään tarkasti, jotta niiden perusteella voidaan tehdä operatiiviset ja taloudelliset suunnitelmat.

Strategian luominen ja toteuttaminen on eri asia. Kuluneina vuosikymmeninä tutkimukset ovat osoittaneet, että suunnitellun strategian ja toteutuneen strategian välillä on kiilu. Palladium Groupin tutkimusten mukaan jopa 90 prosenttia yrityksistä epäonnistuu toteuttaa strategiaansa suunnitelmien mukaan. Syyt strategian epäonnistumiseen vaihtelee, mutta useimmiten se on yksi seuraavista:

### 1. Kommunikointi

- Useissa yrityksissä hyvin pieni osa henkilöstöstä ymmärtää yrityksen strategian täydellisesti

### 2. Palkitsemisen ja kannustimien yhdenmukaisuus

- Usein palkkiot on sidottu lyhyen aikavälin tavoitteeseen. Strategiset päätökset lyhyelle aikajaksolle ei usein ole järkeviä

### 3. Keskittyminen

- Yrityksen johto ei käytä aikaa oleellisiin asioihin, vaan keskittyy pieniin yksityiskohtiin esimerkiksi budjetissa. Tutkimuksen mukaan useissa yrityksissä alle 85% johtajista keskustelelee strategiasta alle tunnin kuukaudessa.

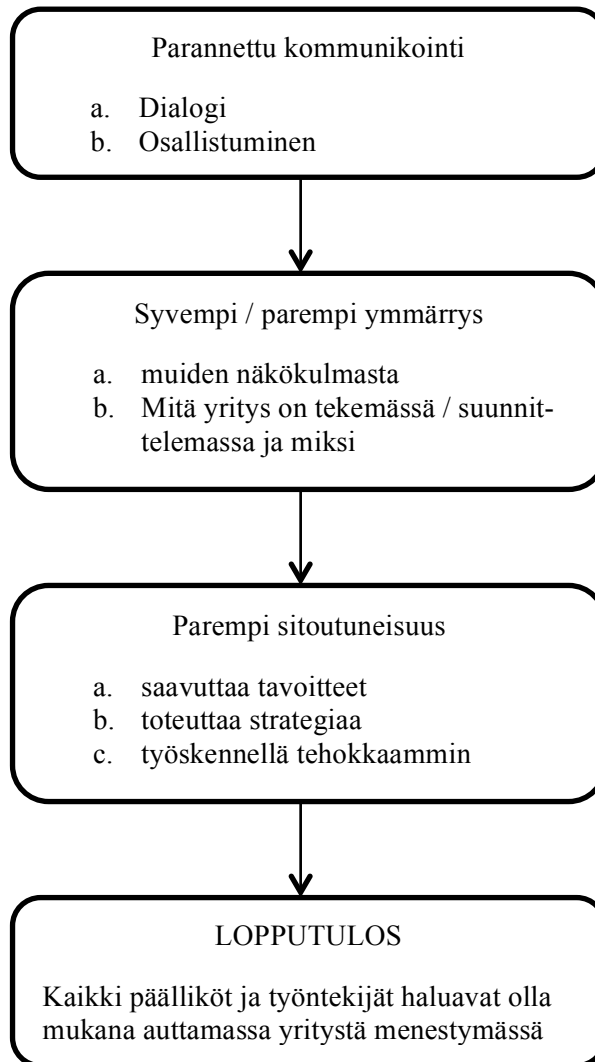
### 4. Resurssit

- Strategian epäonnistuminen on varmaa, ellei strategisia suunnitelmia ole rahoitettu ja resursoitu kunnolla.

(Turban ym. 2014:343-346)

### 5.3 Strategisen johtamisen edut

Strateginen johtaminen mahdollistaa yrityksen olla enemmän ennakoiva kuin reagoiva muokatessaan omaa tulevaisuuttaan. Strategisen johtamisen avulla yritys pystyy muodostamaan itselleen paremman strategian ja tekemään parempia strategisia valintoja.



Kuva 16. Strategisen suunnittelun tuomat edut (David & David 2015: 49)

Avain onnistuneeseen strategiseen johtamiseen on kommunikaatio. Kuvassa 16 kuvataan avoimen keskustelun tuomia hyötyjä. Kun strategiasta keskusteluun osallistuu koko yrityksen henkilöstö, he ymmärtävät strategian paremmin ja sitoutuvat tukemaan yritystä. Kun henkilöstä ja päälliköt ymmärtävät mitä yritys on tekemässä ja miksi, he usein

tuntevat olevansa osa yritystä ja tukevat päätettyä strategiaa. Päälliköistä ja työntekijöistä tulee yllättävän luovia ja innovatiivisia kun he ymmärtävät ja tukevat yrityksen missiota, tavoitteita ja strategiaa. Sitoutunut henkilöstö yleensä ymmärtää yhteyden heidän palkan ja palkkioiden ja yrityksen menestyksen välillä. (David & David 2015: 49)

Strateginen johtaminen on oppimista, auttamista, kouluttautumista ja prosessien tuke-  
mista. Keskustelu strategisesta johtamisesta on tärkeämpää kuin dokumentit strategisesta suunnitelmasta. Huonoimmillaan strategiaa suunnitellaan yksin ja sen jälkeen esitellään valmis strategia johtajille. Hyvien strategisten päätösten tekeminen on omistajien ja johtajien vastuulla, mutta on tärkeää, että päälliköt ja työntekijät ovat mukana muodostamassa, toteuttamassa ja arvioimassa aktiviteetteja. Yhä useammat yritykset käyttävät strategista johtamista tehdessään päätöksiä. Strateginen johtaminen ei ole kuitenkaan ta-  
menestyksestä, siitä voi olla myös haittaa mikäli sitä toteutetaan huonosti. Tutkimukset osoittavat, että yritykset, jotka käyttävät strategisen johtamisen konsepteja ovat onnistuneet parantamaan myyntiä, kannattavuutta ja tuottavuutta merkittävästi verrattuna yrityksiin, jotka eivät systemaattisesti suunnittele aktiviteettejaan. Myös mikäli yrityksen suunnittelujärjestelmä on hyödyntänyt strategisen suunnittelun teorioita yritys on yleensä menestynyt taloudellisesti erittäin hyvin pitkällä tähtäimellä. Menestyneet yritykset pystyvät tekemään tarkempia ennusteita lyhyelle ja pitkälle tähtäimelle. Huonosti menestyvät yritykset eivät onnistu ennusteissaan sekä niiden strategiasta vastaavat henkilöt keskittyvät yleensä ratkaisemaan sisäisiä ongelmia. He monesti aliarvioivat kilpailijansa ja yliarvioivat omat vahvuutensa. (David ym. 2015:49-50)

Strateginen johtaminen tuo yritykselle muitakin kuin taloudellisia etuja, kuten tietoa ulkoisista uhista, parempaa tietoa kilpailijoiden strategiasta, parempi työntekijöiden tuottavuus, vähemmän vastustusta muutoksille sekä parempi ymmärrys menestyspalkitseminen suhteesta. Jotkut yritykset eivät kuitenkaan toteuta strategista suunnittelua. Jotkut yritykset taas toteuttavat, mutta eivät saa siihen tukea päälliköiltä tai työntekijöiltä. Kymmenen yleisintä syytä miksi strategista suunnittelua ei ole tehty tai se on tehty huonosti:

1. Koulutuksen puute
2. Ei ymmärretä tai arvosteta suunnittelun tärkeyttä
3. Ei rahallisia palkitsemisia suunnittelun tekemisestä
4. Ei rangaistuksia suunnittelun puuttumisesta
5. Kiire sammutella tulipaloja
6. Suunnittelua pidetään ajan hukkana

7. Laiskuus. Tehokas suunnittelu vie aikaa – aika on rahaa
8. Nykyisen menestyksen sisältö. Ei ymmärretä, että tämän päivän menestys ei ole tae huomisen menestyksestä
9. Yliluottavaisuus
10. Huonoja kokemuksia strategisesta suunnittelusta joskus tai jossain

Strateginen suunnittelu on sitova, monimutkainen ja monitahoinen prosessi, joka vie yrityksen tuntemattomalle alueelle. Strateginen suunnittelu ei tarjoa valmiita ratkaisuja menestykseen, se tarjoaa yritykselle viitekehysten esittämään kysymyksiä ja ratkaisemaan ongelmia. Menestyäkseen yrityksen on osattava varoa strategisen suunnittelun sudenkuoppia. Alla on listattuna yleisimpiä strategisen suunnittelun sudenkuoppia:

- Strategisella suunnittelulla pyritään saamaan kontrolli päätöksistä ja resursseista
- Strategista suunnittelua tehdään vain akkreditoinnin ja sääntelyn vuoksi
- Siirrytään mission kehittämisestä strategian muodostamiseen
- Epäonnistutaan kertomaan suunnitelmat henkilöstölle
- Ylin johto tekee intuitiivisia päätöksiä jotka ovat ristiriidassa virallisen suunnitelman kanssa
- Ylin johto ei aktiivisesti tue strategisen suunnittelun prosesseja
- Epäonnistutaan suorituksen mittaamisessa
- Delegoidaan suunnitteleminen suunnittelijoille
- Ei saada avainhenkilöitä sitoutumaan suunnittelun jokaiseen vaiheeseen
- Epäonnistutaan muutosta kannustava ilmapiiri
- Pidetään suunnittelua tarpeettomana tai turhana
- Keskitytään nykyisiin ongelmiin liikaa, eikä tehdä kunnollista suunnittelua
- Tehdään suunnittelusta niin muodollista, että joustavuus ja luovuus tukahdutetaan

Jotkut yritykset käyttävät paljon aikaa strategiseen suunnitteluun, mutta sen toteuttaminen epäonnistuu. Muutos ja tulokset eivät tapahdu suunnittelemalla, vaan se vaatii suunniteltuja toimenpiteitä. Strategia vaatii jatkuvaa arviointia, koska liiketoimintaympäristö muuttuu koko ajan ja suunniteltu strategia voi vaatia muutoksia. (David ym. 2015: 50-51)

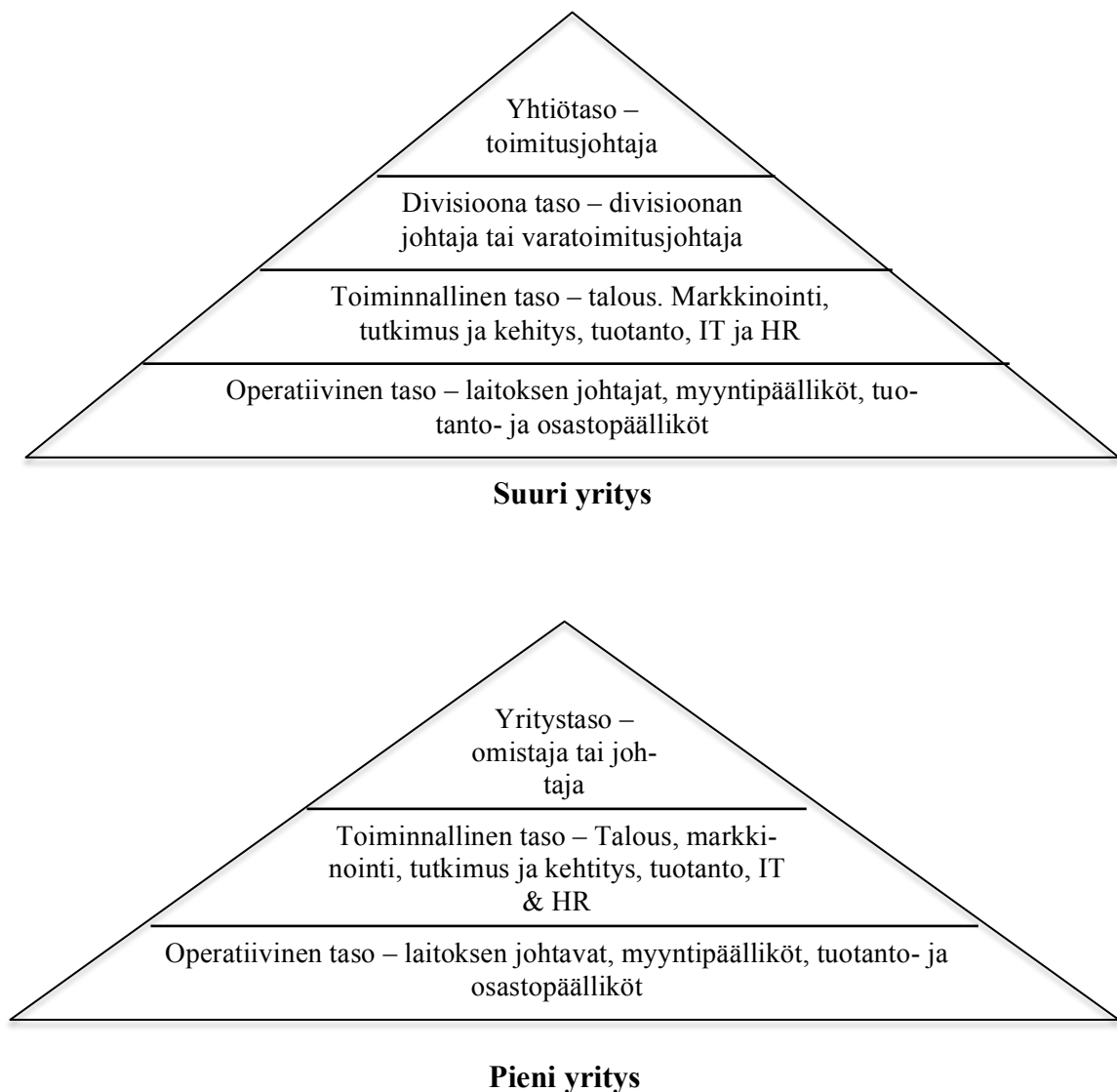
Strategisesta johtamisesta ei saa tulla itseään ylläpitävää byrokraattista mekanismia. Sen tulisi olla oppimisprosessi, joka tutustuttaa päälliköt ja työntekijät strategiaan ja tarjota vaihtoehtoja ongelmien ratkaisemiseksi. Tärkeä ohjenuora tehokkaaseen strategiseen johtamiseen on avoimuus. Jokaisen tulee olla valmis uuteen informaatioon, uusiin näkökulmiin, ideoihin ja mahdollisuuksiin. Taulukossa 6 on listattu 17 ohjetta tehokkaan strategian toteuttamiseksi. Strategiset päätökset vaativat kompromisseja pitkän ja lyhyen strategian välillä tai voiton maksimoinnin ja osakkeenomistajien tahdon välillä. Yleinen virhe on toteuttaa useaa strategiaa samanaikaisesti. Tämä vie paljon yrityksen resursseja ja se samalla vaarantaa kaikki toteutettavat strategiat. (David ym.2015: 51-52)

Taulukko 6. Seitsemäntoista ohjetta tehokkaaseen strategiaan. (David ym. 2015:52)

1. Sen pitää olla ihmisten prosessi eikä paperiprosessi
2. Sen pitää olla oppimisprosessi päälliköille ja työntekijöille
3. Sen pitää olla sanoja tuettuna luvuilla mieluummin kuin lukuja tuettuna sanoilla
4. Sen pitää olla yksinkertainen ja epäritiininomainen
5. Sen tulee vaihdella toimeksiannoista, tiimin jäsenistä, kokouskäytännöistä ja ajankäytöstä
6. Sen tulee haastaa oletukset yrityksen nykyisestä strategiasta
7. Sen tulee käsitellä huonot uutiset
8. Sen tulee olla avoin, ottavainen ja oppiva
9. Se ei saa olla byrokraattinen mekanismi
10. Se ei saa olla rituaalinen ja jäykkä
11. Se ei saa olla liian muodollinen tai ennustettava
12. Se ei saa sisältää ammattislangia tai olla vaikeaselkoinen
13. Se ei saa olla muodollinen kontrollointi järjestelmä
14. Se ei saa olla piittaamaton laadullista informaatiota kohtaan
15. Tekniikasta vastaavat ihmiset eivät saa kontrolloida sitä
16. Älä toteuta liian montaa strategiaa yhtä aikaa
17. Vahvista jatkuvasti ”Hyvä etiikka on hyvää bisnestä” –politiikkaa

## 5.4 Strategia tasot

Strategian toteuttaminen ei ole vain ylimmän johdon tehtävä. Keski- ja alemman johdon tulee olla myös mukana suunnittelemassa strategiaa. Suurissa yrityksissä strategialla on neljä tasoa: yhtiötaso, divisioonataso, toiminnallinen taso ja operatiivinen taso. Kuvassa 17 on kuvattu strategiatasot erikokoisissa yrityksissä.



Kuva 17. Strategiatasoista vastuussa olevat henkilöt (David ym. 2015: 137).

Suurissa yrityksissä strategiasta on vastuussa yhtiötasolla toimitusjohtaja, divisioonatasolla divisioonan johtaja tai varatoimitusjohtaja, toiminnallisella tasolla talousjohtaja tai muun osaston päällikkö ja operatiivisella tasolla strategiasta vastaa yksikönpäälliköt tai aluemyyntipäälliköt. Pienissä yrityksissä ylimmän tason strategiasta vastaa omistaja tai johtaja, ja muuten vastuut jakautuvat samalla tavalla kuin suurissa yrityksissä. (David ym 2015: 137)

### 5.5 Suorituskykymittarit

Yrityksen menestymistä asetettujen strategisten tavoitteiden suhteen täytyy mitata, jotta nähdään tuleeko yritys strategisesti onnistumaan. Esimerkiksi myyntimiehen tekemistä mitattaessa saadaan luku, joka kertoo, että 50% tämän kuukauden kaupoista onnistui. Tämä luku ei yksinään kerro paljosta, mutta jos tiedetään, että saman myyntimiehen onnistumisprosentti viimekuussa oli 30% on myyntimies tässä kuussa onnistunut hyvin. Mikäli tähän lisätään vielä tieto, että myyntimiesten keskimääräinen onnistumisprosentti on 80%, tämä kyseinen myyntimies on pärjännyt viimeaikoina huonosti. (Turban ym. 2014:354)

Suorituskykyä mitataan usein suorituskykymittarilla (KPI, Key Performance Indicator), jonka avulla mitataan yrityksen tekemisiä tavoitteiden suhteen. KPI on moniulotteinen ja sillä on monia ominaisuuksia:

- Strategia
- Tavoitteet
- Alueet
- Enkoodaukset
- Aikaikkunat
- Vertailuarvot

KPI:llä voidaan mitata mennyttä tai tulevaa. Menneen ajan KPI:t liittyvät yleensä talouden mittareihin, mutta ei välttämättä aina. Tulevaisuuden KPI:tä kutsutaan myös operatiiviseksi KPI:ksi. Näillä mitataan yleensä aktiviteetteja joilla on suuri vaikutus menneen ajan suorituskykymittareihin, usein siis yrityksen talouteen liittyviin mittareihin. Operatiivisilla suorituskyky mittareilla mitataan mm. asiakastyytyvää asiakaspalvelua, myyntiä sekä myynnin suunnittelua ja ennusteita. (Turban ym. 2014:354).

## 5.6 Balanced scorecard

Parhaiten tunnettu liiketoiminnan suorituskyvyn johtamisen työkalu on Balanced Scorecard(BSC). Ylätasolla Balanced Scorecard on suorituskyvyn mittari ja johtamismetodologia joka auttaa tulkitsemaan yrityksen taloutta, asiakkaita, sisäisiä prosesseja, oppimista ja kasvua suoritettaviksi toimenpiteiksi. BSC auttaa linjaamaan yrityksen toimenpiteet strategiansa mukaisiksi. Tämä tapahtuu kuudessa vaiheessa:

1. Strategian luominen
2. Strategian suunnittelu
3. Yrityksen strategian yhtenäistäminen
4. Operatiivinen suunnittelu
5. Tarkkailu ja oppiminen
6. Testaaminen ja strategian muokkaaminen

Taulukossa 7 on kuvattu kuvitteellisen yrityksen tasapainotettu tuloskortti. Taulukosta näkee, että yrityksellä on neljä tavoitetta, yksi kutakin tasapainotetun tuloskortin näkökulmaa kohti. (Turban ym. 2014: 360-361)

Taulukko 7. Strategiakartta ja tasapainotettu tuloskortti (Turban ym. 2014: 361).

	<b>Strategiakartta</b>	<b>Balanced Scorecard</b>		<b>Strategiset aloitteet</b>
<b>Talous</b>	Kasvata nettotuloa	Nettotulon kasvatus	Kasvata 25%	
<b>Asiakas</b>	Kasvata asiakasuskollisuutta	Asiakasuskollisuus prosentti	Kasvata 15%	Muuta lisensointia ja ylläpidä sopimuksia
<b>Prosessit</b>	Paranna Call Centerin suorituskykyä	Ongelman käsittelyaika	Paranna 30%	Standardoi call centerin prosessit
<b>Oppiminen ja kasvu</b>	Vähennä työntekijöiden vaihtuvuutta	Vapaaehtoisten vaihtuvuusaste	Vähennä 25%	Palkkojen ja bonusten kasvatus

## 5.7 Six Sigma

Sigma ( $\sigma$ ) on kreikkalainen kirjain, jota on käytetty tilastotieteessä ilmaisemaan vaihtelevuutta. Laatumittauksessa vaihtelevuus tarkoittaa defectien lukumäärää. Six sigma on suorituskyvyn johtamismetodologia, jonka avulla pyritään pienentämään havaittujen defectien määrää lähemmäksi nolaa. Six Sigmaa on käytetty enemmän prosessien kehittämiseen kuin itse suorituskyvynjohtamiseen. Prosessien kehittämismetodologiana Six Sigmaa voi käyttää prosessien valvontaan, ongelmakohtien paikantamiseen ja parannuskeinojen soveltamiseen. Six Sigmassa liiketoiminta nähdään ryppäänä prosesseja. Liiketoimintaprosessi on taas rypäs aktiviteetteja, jotka tekevät toimenpiteitä toisia prosesseja tai ihmisiä varten. (Turban ym. 2014: 362)

Six Sixma perustuu DMAIC suorituskyvyn parannusmalliin. DMAIC tulee sanoista:

- Määrittele (**Define**). Määrittele kehitysaktiviteetin tavoitteet ja rajat.
- Mittaa (**Measure**). Mittaa nykyistä järjestelmää. Luo kvantitatiivisia mittareita, jotka tuottavat statistisesti validia dataa.
- Analysoi (**Analyze**). Analysoi järjestelmä tunnistaaksesi tavat eliminoida erot nykyisessä järjestelmässä tai prosesseissa verrattuna haluttuun tilaan.
- Paranna (**Improve**). Etsi tavat tehdä asiat paremmin, halvemmalla tai nopeammin. Käytä projektinjohto- ja muita menetelmiä apuna.
- Ohjaa(**Control**). Luo järjestelmä, joka varmistaa saavutetun tilan säilymisen parannusprojektin jälkeen.

Six Sigma ei eroa muista liiketoiminta metodologioista. Tässä, niin kuin muissakin malleissa tehdään suunnitelmat ja kehitetään mittarit niiden seuranta varten. Jotta Six Sixma onnistuisi, se vaatii paljon samoja asioita kuin muutkin metodologiat:

- Six Sigma mukautuu liiketoimintastrategiaan
- Six Sigma tukee liiketoimintatavoitteita
- Avainhenkilöt ovat mukana toteutuksessa
- Projektien valintaprosessi perustuu potentiaaliseen lisäarvoon
- Six Sigma osaajia on riittävästi projekteille
- Projekteja johdetaan oikein
- Tiiminvetäjien taitoja korostetaan
- Tuloksia seurataan tarkasti

Six Sixman onnistumistodennäköisyyttä voi kasvattaa yhdistämällä Six Sixman toimintatapoja tasapainotetun tuloskortin tapojen kanssa. Tällöin yrityksen toimintatavat ovat suoraan sidottuna yrityksen strategisiin tavoitteisiin ja päämääriin. (Turban ym. 2014: 362-366)

## 6 YHTEENVETO

Tietovarastointi ja Business Intelligence voidaan toteuttaa monella eri tavalla. Tietovarastosta voidaan tehdä pieni esimerkiksi vain excel-tilukko, joka palvelee vain sen toteuttanutta osastoa. Tietovarastosta voidaan myös tehdä koko organisaation kattava tietovarasto (EDW), johon tuodaan tietoa useista eri operatiivisista järjestelmistä. Pienemmissä tietovarastoissa ei välttämättä ole käytössä tietovarastoa lainkaan, vaan niihin on haettu siihen tarvittu data operatiivisesta järjestelmästä. EDW ratkaisuihin taas käytetään massiivista tietovarastoa, jossa on dataa koko organisaation tasolta. EDW:hen voidaan myös tuoda ulkopuolista tietoa, jolla kuvataan esimerkiksi kilpailijoita tai kilpailuympäristöä.

Business Intelligence toteutukseen on myös monia eri tapoja. Voidaan tehdä yksikkökohtaisia excel-tilukoita, joihin haetaan tietoa operatiivisesta järjestelmästä. Yleisimmin kuitenkin Business Intelligence -ratkaisut toteutetaan laajemmin yritystasolla. Tällöin tieto haetaan koko organisaation kattavasta tietovarastosta. Tietovaraston käyttö Business Intelligence -ratkaisuihin on yleisempää, koska BI vaatii muokattua tietoa, jota ei saada suoraan operatiivisesta järjestelmästä. Tiedon vieminen tietovarastoon tapahtuu ETL-prosessin kautta. ETL-prosessissa tieto muokataan oikeaan muotoon BI-raportteja varten. Monesti IT-osasto tekee tietovarastosta valmiita raportteja, joista eri osastot voivat hakea tietoja omiin tarpeisiinsa.

Tietovarasto ja BI palveluntarjoajia on useita, ja heiltä on mahdollisuus ostaa järjestelmien tarvitsemat laitteistot ja ohjelmistot. Nykyään on myös yleistä, että yritys ostaa tietovarasto ja BI järjestelmän SaaS (Software as a Service, ohjelmisto palveluna) tai PaaS (Platform as a Service, alusta palveluna) mallin mukaisesti. Tällöin yrityksen ei tarvitse huolehtia tietojärjestelmien ylläpidosta tai kehityksestä, vaan maksavat ohjelmistoista käytön mukaan. Vastuu laitteistoista ja tietoverkoista on palvelun tarjoajalla.

Yritys tarvitsee menestyäkseen strategian. Strategia vaatii johdolta ja koko yrityksen henkilöstöltä sitoutumista noudattamaan strategiaa. Strategian tulee olla selkeä ja kaikkien tulee ymmärtää se. Onnistuessaan strategiassa yrityksen johdon tulee tehdä strategisia päätöksiä. Näiden päätösten avuksi yrityksellä on Business Intelligence. BI on johdon työkalu, josta päätöksentekijät saavat tietoa, jotta voisivat tehdä parempia päätöksiä. Näistä järjestelmistä saadun datan avulla yrityksen johdon tulee tehdä strategisia päätöksiä, joiden avulla ohjataan yritys menestykseen. Business Intelligence on siis eräänlainen silmukka. Ensin kerätään tietoa ja analysoidaan sitä. Tämän tiedon perus-

teella tehdään strategisia päätöksiä ja ohjataan yritystä. Nämä päätökset ja toimet vaikuttavat dataan, jota kerätään operatiivisista järjestelmistä. Nyt silmukka lähtee uudelleen alusta ja kerätään taas dataa ja taas näiden tietojen perusteella tehdään uusia strategisia päätöksiä. Tähän silmukkaan kuuluu tietenkin kerätyn tiedon puhdistus, muokkaaminen oikeaan muotoon ja lataaminen tietovarastoon. Vaikka BI on monelle yritykselle kriittinen työkalu, täytyy kuitenkin muistaa, että mikään järjestelmä ei takaa yritykselle menestystä. BI järjestelmät antavat päätöksentekijöille tietoa, joka helpottaa päätöksen tekemistä. Päätöksistä on aina vastuussa ihminen.

BI-työkalujen läpikäymistä tai tietovarastointia ja BI-järjestelmän käyttöönottoa ei tutkielmassa tutkittu. Työkalujen vertaileminen olisi vaatinut kaikkien työkalujen asennuksen ja testidatan syöttämisen järjestelmään. Työkalujen vertailu asennuksineen ja konfigurointineen olisi ollut liian aikaa vievä prosessi tutkimukseen laajuuteen nähden.

Jatkotutkimuksille jäi aihetta. Business Intelligence -raporttien tai valmiiden mittaristojen kehittämistä olisi aiheellista tutkia. Myös tietovarastoinnin ETL-prosesseja sekä reaaliaikaista tietovarastointia olisi aihetta tutkia tarkemmin. Lisäksi koko tietovarastoinnin ja Business Intelligencen käyttöönottoa olisi mielekästä tutkia perusteellisesti.

## LÄHDELUETTELO

- Al-Aqrabi, Hussain, Liu, Lu, Hill, Richard, Antonopolous, Nick (2014). Cloud BI: Future of business intelligence in the Cloud. *Jouranal of Computer Systems and System Sciences [online]* 81: 1, 85-96. Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0022000014001007>
- Bansal, Srividya K., Sebastian Kageman (2015). Integrating Big Data: A Semantic Extract-Transform-Load Framework. *Computer [online]* 48: 3, 42-50. Saatavissa: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7063172&isnumber=7063161>
- Chang, Victor (2014). The Business Intelligence as a Service in the Cloud, *Future Generation Computer Systems [online]* 37: 512-534 Saatavissa: <http://www.sciencedirect.com/science/article/pii/S0167739X13002926>
- David, Fred R., Firest R. David (2015). *Strategic management : concepts and cases*. 15. painos. Essex: Pearson.
- Farooq, Farrah (2013). The data warehouse virtualization framework for operational business intelligence, *Expert Systems [online]* 30: 5, 451-472. Saatavissa: <http://search.ebscohost.com.proxy.tritonia.fi/login.aspx?direct=true&db=buh&AN=91973390&site=ehost-live>
- Gash, David, Thilini Ariyachandra, Mark Frolick (2011). Looking to the Clouds for Business Intelligence, *Journal of Internet Commerce [online]* 10:4, 261-269. Saatavissa: <http://dx.doi.org/10.1080/15332861.2011.622694>
- Giovinazzo, William A. (2000). *Object-oriented data warehouse design: building a star schema*. Upper Saddle River, New Jersey: Prentice Hall.
- Hovi, Ari, Henriikki Hervonen, Heikki Koistinen (2009). *Tietovarastot ja business intelligence*. 1. painos. Porvoo: WS Bookwell.
- Hovi, Ari, Jouni Huotari, Tapio Lahdenmäki (2005). *Tietokantojen suunnittelu & indeksointi*. 1. painos. Porvoo: WS Bookwell.
- Oracle (2013). *Oracle® Business Intelligence Standard Edition One Tutorial*. Saatavissa: [http://docs.oracle.com/html/E10312\\_01/dm\\_concepts.htm](http://docs.oracle.com/html/E10312_01/dm_concepts.htm)

- Salminen Ari (2011). *Mikä on kirjallisuuskatsaus? Johdatus kirjallisuuskatsauksen tyyppeihin ja hallintotieteellisiin sovelluksiin*. Vaasan yliopisto. Vaasan yliopiston julkaisuja. Vaasa.
- Shmueli, Galit, Nitin R. Patel, Peter C. Bruce (2010). *Data Mining for Business Intelligence: concepts, techniques, and applications in Microsoft Office Excel with XLMiner*. 2. painos. Hoboken, N.J.: Wiley.
- Turban, Efraim, Ramesh Sharda & Dursun Delen (2011). *Decision Support and Business Intelligence Systems*. 9. painos. Boston, MA: Pearson.
- Turban, Efraim, Ramesh Sharda, Dursun Delen & David King (2011). *Business intelligence: a managerial approach*. 2. painos. Boston; London: Prentice Hall.
- Turban, Efraim, Ramesh Sharda & Dursun Delen (2014). *Decision Support and Business Intelligence Systems*. 9. painos. Boston, MA: Pearson.
- W3C Recommendation 26 February 2015, *Linked Data Platform 1.0*, <http://www.w3.org/TR/2015/REC-ldp-20150226/#intro>
- Xavier, Cristiano, Moreira Fernando (2013), Agile ETL. *Science Direct [online] 9*: 381-387, Saatavissa: <http://www.sciencedirect.com/science/article/pii/S2212017313001965>