

AVOIMEN LÄHDEKOODIN BUSINESS INTELLIGENCE -RAPORTOINTI

Jarmo Kortelainen

Opinnäytetyö
Toukokuu 2010

Tietojenkäsittely
Luonnontieteiden ala





Tekijä(t) KORTELAINEN, Jarmo	Julkaisun laji Opinnäytetyö	Päivämäärä 10.05.2010
	Sivumäärä 57	Julkaisun kieli Suomi
	Luottamuksellisuus () saakka	Verkojulkaisulupa myönnetty (X)
Työn nimi AVOIMEN LÄHDEKOODIN BUSINESS INTELLIGENCE -RAPORTOINTI		
Koulutusohjelma Tietojenkäsittelyn koulutusohjelma		
Työn ohjaaja(t) IMMONEN, Jarkko		
Toimeksiantaja(t) Intellica Solutions Oy		
Tiivistelmä <p>Opinnäytetyön toimeksiantaja Intellica Solutions Oy on jyväskyläläinen business intelligence -ratkaisujen toimittamiseen erikoistunut yritys. Intellica Solutions on käyttänyt ratkaisuissaan yksinomaan kaupallisia tuotteita. Tutkimuksen tarkoituksena oli selvittää, kuinka hyvin avoimen lähdekoodin business intelligence -raportointivälineet vastaavat ominaisuuksiltaan ja teknologisilta ratkaisuiltaan kaupallisia business intelligence -tuotteita.</p> <p>Opinnäytetyö on tyypiltään tapaustutkimus. Työn teoriaosuus käsittelee business intelligenen taustat ja tavoitteet sekä kuvaa koko business intelligenen arkkitehtuurin ja perustyövälineet. Lisäksi teoriaosuudessa esitellään avoimen lähdekoodin määritelmä sekä tutkimuskohteina olleiden työvälineiden lisenssit. Työn tutkimusosuus jakautui kahteen kokonaisuuteen, varsinaisiin raportointiohjelmiin sekä raporttien jakelujärjestelmiin. Raportointiohjelmissä tutkimukseen valittiin JasperReports/iReport ja Pentaho Report Designer, jakelujärjestelmistä JasperServer sekä Pentaho Business Intelligence Server. Tutkimus toteutettiin testaamalla ja havainnoimalla toimeksiantajan työvälineille asettamia ominaisuusvaatimuksia.</p> <p>Työn tuloksena syntyi yksityiskohtainen raportti avoimen lähdekoodin business intelligence -raportointivälineiden tärkeimmistä ominaisuuksista, teknologisista ratkaisuista ja puutteista. Tutkimuksen tulosten avulla toimeksiantaja sai hyvän käsityksen siitä, kuinka avoimen lähdekoodin työvälineet soveltuvat ominaisuuksiensa puolesta erityyppisiin projekteihin, ja pystyi laajentamaan liiketoimintaansa avoimen lähdekoodin business intelligence -ratkaisuihin.</p>		
Avainsanat (asiasanat) avoin lähdekoodi, business intelligence, raportointi, ohjelmistot		
Muut tiedot		



Author(s) KORTELAINEN, Jarmo	Type of publication Bachelor's Thesis	Date 10052010
	Pages 57	Language Finnish
	Confidential () Until	Permission for web publication (X)
Title OPEN SOURCE BUSINESS INTELLIGENCE REPORTING		
Degree Programme Business Information Systems		
Tutor(s) IMMONEN, Jarkko		
Assigned by Intellica Solutions Oy		
Abstract <p>The thesis was assigned by Intellica Solutions Oy, a Jyväskylä-based company focused on delivering business intelligence solutions. Intellica Solutions have been working exclusively with commercial business intelligence products. The purpose of the study was to determine how well the open source business intelligence reporting tools correspond to the commercial business intelligence products, in terms of software features and technological solutions.</p> <p>The thesis is a case study. The theoretical part deals with the backgrounds and objectives of business intelligence, and describes the business intelligence architecture and basic business intelligence tools. In addition, the theoretical part presents the open source definition as well as software licenses for the tools that were the subjects of the study. The empirical part of the thesis is divided into two parts, the actual reporting tools and reporting servers. The selected reporting tools were JasperReports/iReport and Pentaho Report Designer, and selected reporting servers were JasperServer and Pentaho Business Intelligence Server. The study was conducted by testing and observing the selected tools for feature requirements set by Intellica Solutions.</p> <p>The study resulted in a detailed report on key features, technological solutions and absences of open source business intelligence reporting tools. By the means of the results, Intellica Solutions gained a good impression on how open source reporting tools were suitable for different types of business intelligence projects by their features, and were able to expand its business to open source business intelligence solutions.</p>		
Keywords open source, business intelligence, reporting, software		
Miscellaneous		

SISÄLTÖ

KÄSITTEISTÖ	3
1 JOHDANTO	4
2 TUTKIMUSASETELMA.....	6
2.1 TUTKIMUKSEN TOIMEKSIANTAJA, TAVOITTEET JA RAJAUKSET	6
2.2 TUTKIMUSMENETELMÄT	7
2.3 TUTKIMUSKYSYMYKSET.....	8
3 BUSINESS INTELLIGENCE	9
3.1 HISTORIA JA KEHITYS	9
3.2 MÄÄRITELMÄ.....	10
3.3 TAVOITTEET.....	10
3.4 PERUSARKKITEHTUURI.....	12
3.5 RAPORTOINTI.....	14
3.5.1 <i>Raporttien luominen</i>	14
3.5.2 <i>Raporttien jakelu</i>	17
3.6 MUUT PERUSTYÖKALUT	18
3.6.1 <i>Mittaristot</i>	18
3.6.2 <i>OLAP</i>	19
3.6.3 <i>Tiedon louhinta</i>	19
4 AVOIN LÄHDEKODI.....	21
4.1 HISTORIA	21
4.2 MÄÄRITELMÄ.....	22
4.3 LISENSIT	24
4.3.1 <i>GNU General Public License</i>	24
4.3.2 <i>GNU Affero General Public License</i>	25
4.3.3 <i>GNU Lesser General Public License</i>	25
4.3.4 <i>Mozilla Public License</i>	25
5 TUTKIMUKSEN TOTEUTUS	27
5.1 JASPERREPORTS	28
5.2 PENTAHO	29
6 TUTKIMUKSEN TULOKSET	30
6.1 RAPORTOINTIOHJELMAT: JASPERREPORTS / IREPORT JA PENTAHO REPORT DESIGNER	30
6.1.1 <i>Esittely</i>	30
6.1.2 <i>Asennus</i>	31
6.1.3 <i>Datayhteydet</i>	31
6.1.4 <i>Tietolähteiden käsittely</i>	33
6.1.5 <i>Raporttien rakenne</i>	35
6.1.6 <i>Parametrit</i>	37
6.1.7 <i>Kaavat</i>	38
6.1.8 <i>Tiedon ryhmittely ja lajittelu</i>	39
6.1.9 <i>Kaaviot</i>	39
6.1.10 <i>Ristiintaulukointi</i>	40
6.1.11 <i>Aliraportit</i>	42
6.1.12 <i>Raporttien tallennus</i>	43
6.2 JAKELUJÄRJESTELMÄT: JASPERSERVER JA PENTAHO BUSINESS INTELLIGENCE SERVER....	45
6.2.1 <i>Asennus</i>	45
6.2.2 <i>Raporttien asennus</i>	46
6.2.3 <i>Raporttien ajaminen</i>	47
6.2.4 <i>Raporttien ajastus</i>	47
6.2.5 <i>Raporttien katselu</i>	48
6.2.6 <i>Käyttäjän todentaminen</i>	49
6.2.7 <i>Käyttöoikeudet</i>	49

6.3 DOKUMENTAATIO.....	51
7 JOHTOPÄÄTÖKSET JA POHDINTA	52
LÄHTEET	55

KUVIOT

KUVIO 1. Business intelligence -ratkaisun perusarkkitehtuuri	12
KUVIO 2. Tietokantayhteyden luominen iReportissa.....	32
KUVIO 3. iReportin Services-paneeli	33
KUVIO 5. Raportin perusrakenne JasperReportsissa	36
KUVIO 8. Kaavion lähdetietojen määrittäminen Pentahossa	40
KUVIO 9. Parametrin liittäminen aliraporttiin iReportissa.....	42
KUVIO 10. Pentahon raporttia käännettäessä kaavion selite leikkaantuu.	44
KUVIO 11. Raporttien ajastaminen JasperServerissä.....	48
KUVIO 12. Käyttöoikeuksien määrittäminen Pentahossa.....	50

KÄSITTEISTÖ

CAS (Central Authentication Service) on alunperin Yalen yliopistossa luotu avoin kertakirjautumispalvelu. Kertakirjautuminen mahdollistaa useiden verkkosovellusten tai -palvelujen käytön yhdellä kirjautumisella.

JAAS (Java Authentication and Authorization Service) on ohjelmointirajapinta käyttäjien todentamiseen ja käyttöoikeuksen määrittelyyn.

Java on laitteistoriippumaton olio-ohjelmointikieli.

JDBC (Java Database Connectivity) on järjestelmäriippumaton ohjelmointirajapinta tietokantojen käyttämiseksi Java-ohjelmassa.

LDAP (Lightweight Directory Access Protocol) on TCP/IP-verkkoon tarkoitettu hakemistopalveluprotokolla.

MySQL on myös avoimen lähdekoodin lisenssillä saatava tietokannan hallintajärjestelmä.

ODBC (Open Database Connectivity) on Microsoftin kehittämä ohjelmointirajapinta tietokantojen käyttämiseksi.

SQL (Structured Query Language) on standardoitu relaatiotietokantojen kyselykieli.

XML (eXtensible Markup Language) on rakenteisen tiedon esittämiseen suunniteltu merkkäuskieli.

1 JOHDANTO

Kaiken digitaalisen tiedon määrän odotetaan viisinkertaistuvan vuodesta 2008 vuoteen 2012 (Gantz & Reinsel 2009). Yritykset keräävät runsaasti tietoa useista eri sisäisistä ja ulkoisista lähteistä. Tämän talteen kerätyn tiedon tehokas jalostaminen, analysoiminen ja hyödyntäminen ovat eilinehto kilpailussa mukana pysymiselle. Business intelligence -sovellusten avulla yritykset ja organisaatio pääsevät käsiksi toimintansa kannalta tärkeään tietoon. Tutkimusyhtiö Gartnerin (2009) yli 1500 tietohallintopäällikköä ympäri maailmaa kattanut tutkimus osoitti business intelligence -sovellusten olevan yritysten tärkein teknologiainvestointi jo neljättä vuotta peräkkäin (Moore 2009).

Kaupallisten business intelligence -tuotteiden lisensointikustannukset saattavat nousta jo pienilläkin käyttäjämäärillä monelle yritykselle niin korkeiksi, että kaupalliset tuotteet eivät tule kysymykseen business intelligence -investointeja suunniteltaessa. Kaupallisten business intelligence -tuotteiden rinnalle on vasta aivan viime vuosina noussut kourallinen mielenkiintoisia avoimen lähdekoodin vaihtoehtoja. Gartnerin (2008) tutkimuksen mukaan avoimen lähdekoodin business intelligence -ratkaisujen omaksumisen ensimmäinen "early adopter" -vaihe on jo takana. Samalla Gardner arvioi, että vielä vuoteen 2012 asti avoimen lähdekoodin business intelligence -ratkaisuja omaksuvat käyttöönsä lähinnä keskisuuret yritykset. (Bitterer 2008.)

Opinnäytetyössä selvitetään vastaavatko laadukkaimmat olemassa olevat avoimen lähdekoodin business intelligence -raportointivälineet ominaisuuksiltaan kaupallisia tuotteita. Avoimen lähdekoodin raportointivälineistä tutkitaan sellaisia ominaisuuksia, joita on pystynyt

hyödyntämään kaupallisissa tuotteissa jo pitkään ja jotka mahdollistavat raportointivälineiden soveltamisen erityyppisissä projekteissa.

Luvussa 2 kuvataan tutkimusasetelma, jossa esitellään opinnäytetyön toimeksiantaja ja käydään läpi työn tavoitteet, rajaukset, tutkimuksessa käytetyt tutkimusmenetelmät sekä tutkimuskysymykset joihin tutkimuksella pyritään vastaamaan. Tutkimuksen teoreettinen viitekehys jakautuu kahteen kokonaisuuteen. Luku 3 käsittää business intelligenen historian, määritelmän ja tavoitteet sekä esittelee business intelligenen perusarkkitehtuurin, business intelligence -raportoinnin peruspiirteet ja business intelligence -sovellusten muut perustyökalut. Luvussa 4 esitellään avoimen lähdekoodin historia ja määritelmä sekä tutkimuskohteina olevien sovellusten lisenssit. Luvussa 5 kuvataan tutkimuksen toteutus ja esitellään yhtiöt tutkimuskohteina olevien sovellusten takana. Luvussa 6 käsitellään tutkimuksen tulokset ja luvussa 7 tutkimuksen johtopäätökset, pohdinta ja kehitysehdotukset.

2 TUTKIMUSASETELMA

Tutkimusasetelma sisältää toimeksiantajan esittelyn, tutkimukselle asetetut tavoitteet ja rajaukset. Luvussa selvitetään lisäksi tutkimuksen teossa käytettyjä tutkimusmenetelmiä ja esitetään tutkimuskysymykset, joihin tutkimuksella pyritään vastaamaan.

2.1 Tutkimuksen toimeksiantaja, tavoitteet ja rajaukset

Opinnäytetyön toimeksiantaja Intellica Solutions Oy on jyvaskyläläinen business intelligence -ratkaisujen toimittamiseen erikoistunut yritys. Intellica Solutions toimii alkujaan ranskalaisen Business Objects -ohjelmistoyhtiön partnerina ja käyttää toimittamissaan business intelligence -ratkaisuissa kaupallisia SAP Business Objects -tuoteperheen tuotteita. Intellica Solutions toivoi selvitystä avoimen lähdekoodin business intelligence -tuotteiden ominaisuuksista ja niiden soveltuvuudesta eri projekteihin. Intellica Solutions on nähnyt avoimen lähdekoodin business intelligence -tuotteet mahdollisuutena laajentaa yrityksen liiketoimintaa. Tutkimuksen avulla toimeksiantaja näkee mihin avoimen lähdekoodin -tuotteet pystyvät ja kykenee esittämään asiakkailleen vaihtoehtoisia avoimen lähdekoodin business intelligence -ratkaisuja, mikäli kaupallisten tuotteiden lisenssointikustannukset nousevat liian korkeiksi.

Monipuolisimmat business intelligence -tuotteet ovat erittäin laajoja ohjelmistokokonaisuuksia, eikä niiden riittävän laaja ja yksityiskohtainen tutkiminen ole mahdollista yhdessä opinnäytetyössä. Opinnäytetyön aihe rajattiin koskemaan avoimen lähdekoodin business intelligence -tuotteiden raportointi- ja jakeluominaisuuksia ja -teknologioita. Raportointi on business intelligence -tuotteiden keskeisin ja loppukäyttäjien kannalta kaikkein

oleellisin työväline. Tämän lisäksi raportointi- ja jakelutuotteet ovat business intelligence -ohjelmistojen lisenssointikustannuksissa kaikkein merkittävimmissä osassa.

Tavoitteena on kirjallisuuden ja muiden lähteiden avulla selvittää business intelligencen tavoitteet sekä business intelligence -tuotteiden yleisimpien työvälineiden perusominaisuudet. Työn teoriaosuudessa tutustutaan lisäksi avoimen lähdekoodin määritelmään sekä tutkimuskohteiksi valittujen sovellusten avoimen lähdekoodin lisensseihin. Teoriatietojen ja toimeksiantajan ohjeistuksen perusteella ryhdytään tutkimaan ensin pintapuolisesti erilaisia avoimen lähdekoodin business intelligence -tuotteita, joista lopulta valitaan kaksi raportointisovellusta ja kaksi jakelujärjestelmää tarkempaan tutkimukseen.

Toimeksiantaja on asettanut tutkittaville tuotteille erittäin tarkan listan ominaisuusvaatimuksista, jotka ovat peräisin heidän käyttämistään kaupallisista tuotteista. Tutkimuksessa pyritään ensisijaisesti selvittämään kuinka hyvin valitut tuotteet vastaavat toimeksiantajan esittämiin ominaisuusvaatimuksiin ja samalla arvioidaan valittujen raportointisovellusten ja jakelujärjestelmien monipuolisuutta.

2.2 Tutkimusmenetelmät

Tutkimusmenetelmänä käytetään tapaustutkimusta ja tutkimusotteena on kvalitatiivinen eli laadullinen tutkimus. Kvalitatiivisen tutkimuksen pyrkimyksenä on kohteen mahdollisimman kokonaisvaltainen tutkiminen ja lähtökohtana on aineiston monitahoinen ja yksityiskohtainen tarkastelu. Tutkimuksen kohteet valitaan tarkoituksenmukaisesti ja tutkittavan tiedon hankinnassa luotetaan tutkijan omiin havaintoihin. Havainnoinnin lisäksi aineiston hankinnassa hyödynnetään erilaisia dokumentteja. (Hirsjärvi,

Remes & Sajavaara 2009, 161, 165.)

2.3 Tutkimuskysymykset

Tutkimuksen avulla pyritään vastaamaan seuraaviin tutkimuskysymyksiin:

1. Miten business intelligence -ratkaisuilla voidaan tukea yrityksen liiketoimintaa?
2. Kuinka monipuolisia ja kattavia avoimen lähdekoodin tuotteet ovat raportointiominaisuuksiltaan?
3. Ovatko avoimen lähdekoodin business intelligence -tuotteet ominaisuuksiltaan ja teknologisilta ratkaisuiltaan valmiita korvaamaan kaupalliset tuotteet?

3 BUSINESS INTELLIGENCE

Luvussa selvennetään business intellingencen eri määritelmiä ja muodostetaan lyhyt katsaus business intelligenen historiaan. Luvussa esitellään business intelligence -ratkaisujen perustavoitteet sekä tyypillisin business intelligence -ratkaisun arkkitehtuuri. Lopuksi käydään läpi business intelligence -ratkaisujen keskeisimmät työkalut, niiden käyttötarkoitukset ja perusominaisuudet.

3.1 Historia ja kehitys

Yritysten päätöstekijät ovat hyödyntäneet erilaisia päätöksentekoa tukevia tietojärjestelmiä on jo 60–70 luvuilta asti. Tuolloin puhuttiin johtamisen tietojärjestelmistä (Management Information Systems). 80-luvulla tulivat toimeenpanevan johdon tieto- ja tukijärjestelmät (Executive Information Systems ja Executive Support Systems) sekä ryhmäpäätöksenteon tukijärjestelmät (Group Decision Support Systems). Teknologiyhtiö IBM:n tietojärjestelmätieteilijä Hans Peter Luhn oli kirjoittanut jo vuonna 1958 älykkäistä järjestelmistä artikkelin otsikolla ”A Business Intelligence Systems” (Grimes 2008). Termin Business Intelligence kuitenkin yleistä Howard Dresner vasta vuonna 1989 (Power 2007).

90-luvulla tulivat ensimmäiset liiketoimintaa tukevat tietovarastointiratkaisut, ja internetin yleistyessä nähtiin ensimmäiset web-pohjaiset järjestelmät (Power 2007). Viime vuosina business intelligence -markkinat ovat kasvaneet vauhdilla. Business intelligence - ja analytiikkaohjelmistojen markkinat olivat vuonna 2008 noin 8,8 miljardia Yhdysvaltain dollaria, ja kasvua edellisvuoteen oli lähes 22 % (Pettey & Stevens 2009).

3.2 Määritelmä

Business intelligence ei ole yksiselitteisesti määriteltävissä oleva termi vaan määritelmiä on niin monta kuin on määrittelijääkin. Suppeimmillaan business intelligence nähdään pelkästään tiedon raportointina ja esittämisenä (Azvine, Cui & Nauck 2005). Laajemman määritelmän mukaan business intelligence on koko organisaation laajuinen strategia, joka tukee raportointia, analysointia ja päätöksentekoa organisaation eri tasoilla (Abukari & Jog 2003).

Business intelligence on systemaattinen prosessi tiedon keräämiseen ja analysointiin organisaation kilpailuympäristöstä (Thomas Jr. 2001). Se on datan hankkimista, ymmärtämistä, analysointia ja muuttamista organisaation toimintaa ohjaavaksi tiedoksi liiketoiminnan suorituskyvyn parantamiseksi (Azvine, Cui & Nauck 2005).

Business intelligencellä pyritään parantamaan organisaation strategiaa, taktisia ja operationaalisia päätöksiä hyödyntämällä useista eri lähteistä hankittua dataa, informaatiota ja tietämystä (Thierauf 2001). Abukari ja Jog (2003) tiivistävät business intelligencen niin, että kyseessä on oikean informaation välittämistä oikeille päätöksentekijöille oikeaan aikaan.

3.3 Tavoitteet

Business intelligence -ratkaisujen päätavoitteena on **tukea organisaation päätöksentekoa**. Vaikka järjestelmät ovat muuttuneet valtavasti sitten 60-luvun varhaisimpien päätöksenteon tukijärjestelmien, tavoite on edelleen pysynyt samana. Toimivilla business intelligence -ratkaisuilla voidaan nopeuttaa ja parantaa organisaation kykyä tehdä oikeampia päätöksiä operatiiviselta tasolta aina korkeimpaan johtoon asti. (Hovi, Hervonen & Koistinen 2009, 80.)

Toinen business intelligenen keskeisimpiä tavoitteita on **käyttäjien tietotarpeisiin vastaaminen**. Järjestelmien tavoitteena on paitsi tuottaa informaatiota valmiiksi suunnitellulla tavalla, ennen kaikkea mahdollistaa ennalta määrittelemättömiin tietotarpeisiin vastaaminen nopeasti ja tehokkaasti. Business intelligence -ratkaisujen kehittyneillä kysely- ja raportointiominaisuuksilla voidaan suorittaa ennalta määrittelemättömiä hakuja ja muodostaa tuloksista halutunlaisia raportteja. Tietotarpeisiin liittyen business intelligence -ratkaisujen tavoitteena on myös **parantaa käyttäjien omatoimisuutta** niin, että loppukäyttäjät pystyvät itsenäisesti hakemaan tarvitsemansa tiedot. Tämän tavoitteen täyttymisen edellytyksenä on business intelligence -ratkaisujen helppokäyttöisyys. (Hovi, Hervonen & Koistinen 2009, 80–81.)

Yksi haasteellisimpia business intelligenen tavoitteita on koko **organisaation strategian tukeminen**. Parhaat edellytykset tavoitteen toteutumiseen ovat silloin, kun organisaatiossa on määritelty mitattavat strategiset tavoitteet. Business intelligence -järjestelmien mittaristoratkaisuilla voidaan seurata määritettyjen tunnuslukujen kehittymistä organisaation eri toiminnoissa. (Hovi, Hervonen & Koistinen 2009, 80.)

Näiden tavoitteiden lisäksi business intelligence -ratkaisuilla pyritään tunnistamaan uhkia ja mahdollisuuksia, löytämään organisaation haavoittuvaisimmat kohdat, välttämään yllätyksiä, parantamaan reagointiainiaa ja suojaamaan organisaatiolle arvokasta aineetonta pääomaa (Thomas Jr. 2001, 48).

3.4 Perusarkkitehtuuri

Business intelligence -ratkaisujen arkkitehtuuri rakentuu tavallisesti neljästä eri tasosta (ks. kuvio 1): tietolähteet, integrointi, tietovarasto ja analyysi.

Business intelligence -ratkaisujen tietolähteinä toimivat organisaatioilla tyypillisesti käytössä olevat useat erilaiset operatiiviset tietojärjestelmät, kuten toiminnanohjaus-, asiakkuudenhallinta- ja laskutusjärjestelmät. Myös eri ohjelmistoista saatavia siirtotiedostoja voidaan hyödyntää tietolähteinä.

(Hovi, Hervonen & Koistinen 2009, 86.)



KUVIO 1. Business intelligence -ratkaisun perusarkkitehtuuri (Business Intelligence 2009).

Tiedon integrointi -tasolla tietolähteistä saatava data haetaan ja muokataan yhdenmukaiseksi. Vaiheesta voidaan käyttää myös nimitystä ETL-vaihe, koska integrointi suoritetaan usein ETL-välineillä. ETL-lyhenne tulee sanoista Extract, Transform ja Load eli poiminta, muokkaus ja lataus. ETL:n ensimmäisessä vaiheessa nimensä mukaisesti poimitaan operatiivisten tietojärjestelmien omista tietokannoista halutut tiedot. (Hovi, Hervonen & Koistinen 2009, 86.)

Kun tietojen poiminta on suoritettu siirrytään muokkausvaiheeseen.

Operatiivisiin järjestelmiin on tiedot syötetty usein käsin, joten

muokkausvaiheessa on syytä tarkistaa kaikki poimitut tiedot. Voidaan esimerkiksi etsiä tuplarivejä, tarkistaa että päivämäärät, puhelinnumerot, henkilötunnukset ja vastaavat tiedot ovat halutussa muodossa ja että pakolliseksi määrätyt sarakkeet on täytetty ja mahdollisissa ennalta määrätyissä raja-arvoissa on pysytty. Virheellisiksi havaitut tiedot voidaan joko hylätä kokonaan, kirjoittaa tiedot erillisiin virhetauluihin tai ottaa tiedot virheellisinäkin mukaan, mutta samalla merkitä ne virheellisiksi. (Hovi, Hervonen & Koistinen 2009, 56.)

Vasta tietojen tarkistuksen jälkeen suoritetaan tietojen varsinainen muokkaus. Eri tietolähteistä poimittuja tietoja pyritään yhdistelemään ja yhdenmukaistamaan. Joitain tietoja voidaan pystyä täydentämään olemassa olevien tietojen perusteella ja yhdistelmätietoja voidaan purkaa pienempiin osiin. Muokkauksessa voidaan jo suorittaa ensimmäisiä laskutoimituksia ja lisätä näiden tulokset tietoihin. (Hovi, Hervonen & Koistinen 2009, 56–57.)

ETL:n viimeisessä vaiheessa muokatut tiedot ladataan tietovarastoon. Tietojen lataus voidaan toteuttaa joko ETL-välineillä, tietokantatuotteiden latausohjelmilla tai SQL-kielen insert-lauseilla. Koko integrointi tai ETL-vaihe toteutetaan yleensä ajastetusti ilta- tai yöaikaan, jolloin käyttäjiä on mahdollisimman vähän ja käytettäviä laitetehoja mahdollisimman paljon. ETL-välineet tallentavat lokitietoja, joista voidaan tarkistaa esimerkiksi mitä kaikkea ajon aikana tehtiin ja kuinka kauan toimenpide kesti. (Hovi, Hervonen & Koistinen 2009, 58.)

Kolmannella tasolla tarkistetut ja muokatut tiedot ladataan tietovarastoon. Tietovarasto on paikka, johon eri järjestelmistä poimittu tieto kerätään ja tallennetaan ja josta sitä tarjotaan käyttäjien raportointia ja analysointia varten. Tietovarastot suunnitellaan ja optimoidaan palvelemaan raportointia ja tietojen analyysia, joten sen käyttö ja hakujen tekeminen on huomattavasti

nopeampaa kuin jos käytettäisiin useita erilaisia operatiivisten järjestelmien tietokantoja samanaikaisesti. Näin ei myöskään kuormiteta operatiivisten järjestelmien omia tietokantoja toistuvilla kyselyillä. (Howson 2007, 28–30.)

Arkkitehtuurin viimeinen taso on loppukäyttäjän kannalta näkyvin ja konkreettisin. Viimeisellä tasolla toimivat varsinaiset business intelligence -ohjelmistot. Tällä tasolla käyttäjät voivat esimerkiksi lukea tiedoista koostettuja raportteja, suorittaa hakuja, tietojen analysointia ja louhintaa sekä tarkastella tunnuslukuja. (Hovi, Hervonen & Koistinen 2009, 86.)

3.5 Raportointi

Kuten business intelligenen määritelmässä käy ilmi niin suppeimmillaan koko business intelligence ymmärretään juuri tiedon raportointina. Raportointia pidetäänkin tärkeimpänä business intelligence -ratkaisujen tiedon hyödyntämismuotona.

3.5.1 Raporttien luominen

Business intelligence -ohjelmistoissa raportit rakennetaan tietovarastoon tehtävillä tietokantakyselyillä. Kehittyneimmissä business intelligence -ohjelmistoissa saattaa olla tuki kymmenille eri tietokantajärjestelmille. Yleisimpiä business intelligence -ohjelmistojen käyttämiä tietokantajärjestelmiä ovat Oracle, Microsoft SQL Server, IBM DB2 sekä avoimen lähdekoodin tietokannat MySQL ja PostgreSQL. Tietokannat eivät ole kuitenkaan ainoita raporttien tietolähteitä. Esimerkiksi kaupallisessa Crystal Reports -raportointiohjelmassa tietolähteinä voidaan käyttää tietokantojen lisäksi XML-dokumentteja, palvelimen lokitietoja tai vaikkapa Microsoft Exchange -ryhmäviestiohjelmaa. (Peck 2008, 437, 491.)

Käyttäjille voi tulla tarve saada käyttöönsä ennalta määrittelemättömiä, spontaaneja raportteja (ns. ad hoc -raportit). Pitempiaikaisen business intelligence -ohjelmiston käytön jälkeen saatetaan huomata, että jotkin ad hoc -kyselyistä toistetaan aika ajoin. Tällöin niistä on hyvä tehdä valmiita vakioraportteja. Osa ad hoc -kyselyistä voi kuitenkin olla niin spontaaneja, että tietty kysely tehdään vain sen yhden ainoan kerran. Raportointityökalun tulisi olla helppokäyttöinen, jotta käyttäjät pystyvät tehokkaasti ja nopeasti luomaan ad hoc -kyselyitä organisaation päätöksentekoa tukemaan. Jos ad hoc -kyselyiden laadinta on hankalasti toteutettu, voidaan organisaatiossa joutua tilanteeseen, jossa ainoastaan it-asiantuntijat ja vastaavat henkilöt kykenevät luomaan ad hoc -kyselyjä. Tällöin kyselyjen spontaanius häviää hyvin nopeasti, eikä päätöksentekoa pystytä tukemaan parhaalla mahdollisella tavalla. (Howson 2007, 35–36.)

Laadukkaissa ohjelmistoissa loppukäyttäjien ei tarvitse syöttää varsinaisia kyselylauseita vaan kyselyt pystytään luomaan sovelluksen käyttöliittymässä. Käyttöliittymässä käyttäjä näkee käytettävissä olevat tietoelementit, ulottuvuudet ja mittaritiedot selkokieლისinä ja pystyy hiirellä valitsemaan niistä raportin muodostamisessa tarvittavat tiedot. Kyselyt voidaan myös rajata useilla eri tavoilla koskemaan esimerkiksi vain tiettyä tuotetta vai vuotta. (Hovi, Hervonen & Koistinen 2009, 87–88.)

Yksi raportti voi helposti sisältää satojen asiakkaiden tietoja tai tuhansia tilauksia ja tuotteita. Tästä syystä raporttien ryhmittely ja lajittelu ovat hyvin oleellisia raportoinnin perusominaisuuksia. Esimerkiksi tuotteet tulee voida ryhmitellä tuotemerkkien ja mallien mukaan, asiakkaat lajitella aakkosjärjestykseen ja tilaukset tilauspäivämäärien mukaan. Järkevästi ryhmitellyistä ja lajitelluista raporteista käyttäjän on helpompi löytää etsimänsä tiedot. (Peck 2008, 38–40.) Muita raportoinnin perusominaisuuksia ovat muun muassa erilaiset laskentakaavat ja graafit. Yksinkertaisimmillaan

laskentakaavoja voidaan käyttää esimerkiksi tilausten kokonaissummien laskemiseen, mutta laskentakaavat mahdollistavat myös huomattavasti monimutkaisempien laskutoimitusten tekemisen. Laskentakaavoissa voidaan käyttää esimerkiksi aritmeettisia- ja loogisia operaatioita sekä funktioita. Aritmeettisiä operaatioita ovat perinteiset yhteen-, vähennys-, kerto ja jakolaskuoperaatiot, kun taas loogisia operaatioita ovat esimerkiksi ei-, ja-, tai-, joko-tai-operaatiot. Sovelluksessa rakennetut laskentakaavat voidaan tallentaa ja hyödyntää myöhemmin useissa eri raporteissa. Raporttiin liitettävillä graafeilla pystytään havainnollistamaan raportin pääkohtia. (Bouman & van Dongen 2009, 395–397.)

Raportteihin voidaan tavallisesti liittää myös muuttuvia parametrejä. Parametrien avulla raporttien loppukäyttäjä pystyy helposti asettamaan raportille rajauksia ja hakuehtoja. Parametreiksi voidaan määritellä esimerkiksi tuoteryhmä tai aika. Näitä parametrejä muuttamalla käyttäjä pystyy päivittämään valmista raporttipohjaa vastaamaan paremmin omia tarpeitaan. Hyvät business intelligence -ohjelmistot mahdollistavat parametrien liittämisen suoraan käyttäjätunnuksiin niin, että käyttäjätunnuksen perusteella käyttäjälle annetaan ennalta määritellyt parametrit valmiiksi käyttöön. (Hovi, Hervonen & Koistinen 2009, 90.)

Eri yhtiöiden business intelligence -ohjelmistot tallentavat raportit omiin ohjelmistokohtaisiin tiedostoformaateihin. Ohjelmistojen tiedostoformaatit ovat yleensä yhteensopivia vain samaan tuoteperheeseen kuuluvien ohjelmistojen kanssa. Näin ollen raporttien jakaminen ohjelmiston omassa tiedostoformaattissa ei yleensä ole kannattavaa. Business intelligence -ohjelmistot eivät kuitenkaan rajoita käyttäjää käyttämään pelkästään ohjelmiston omia formaatteja vaan tasokkaista ohjelmistoista voidaan viedä raportteja useammassa eri tiedostoformaattissa. Tyypillisimmin raportteja voidaan viedä PDF (Portable Document Format), RTF (Rich Text Format), XLS

(Excel Spreadsheet) tai HTML (Hypertext Markup Language) -muodoissa, jolloin raporttien lukeminen onnistuu helposti esimerkiksi Adobe Acrobat, Microsoft Excel ja Microsoft Word -ohjelmissa tai suoraan web-selaimessa. (Peck 2008, 419.)

3.5.2 Raporttien jakelu

Raporttien ajaminen voidaan toteuttaa muutamalla eri tavalla. Raportit voidaan ajaa välittömästi, jolloin käyttäjä joutuu jäämään odottamaan tuloksia. Monipuolisten ja laajojen raporttien ajaminen ja odottaminen saattaa kestää hyvinkin kauan. Toinen vaihtoehto on suorittaa raporttien ajaminen taustalla. Sen sijaan, että käyttäjä joutuisi vain jäämään odottamaan tuloksia, pystyy hän sillä aikaa selaamaan esimerkiksi muita, jo valmiita raportteja. Tulosten saavuttua taustalla ajatut raportit tallennetaan ja raportit ovat käyttäjän luettavissa silloin, kun hän haluaa. (Bouman & van Dongen 2009, 411.) Loppukäyttäjät tarvitsevat kuitenkin erilaisia vakioraportteja eli raportteja, jotka luodaan aina tietyin väliajoin, kuten päivittäin tai viikoittain. Monet business intelligence -ohjelmistot mahdollistavat raporttien ajastamisen niin, että raportit voidaan ajaa automaattisesti ennalta määritettyinä aikoina. Viikoittain tarvittavien vakioraporttien ajastaminen voidaan toteuttaa esimerkiksi niin, että raportit ajetaan joka maanantai kello kahdeksan. (Hovi, Hervonen & Koistinen 2009, 165.)

Monet business intelligence -ohjelmistot mahdollistavat myös ajastettujen raporttien lähettämisen sähköpostin välityksellä. Raporttien lähettäminen voidaan automatisoida esimerkiksi niin, että esimiehet saavat omaa osastoaan tai vastuualuettaan koskevat raportit suoraan sähköpostin liitetiedostona tai html-sähköpostina. Samaan tapaan voidaan yrityksen asiakkaille lähettää vaikkapa maksumuistutuksia. (Bouman & van Dongen 2009, 430, 438.)

Business intelligence -ohjelmistojen käyttäjien hallintaan ja todentamiseen on olemassa muutamia eri vaihtoehtoja. Käytännössä jokainen business intelligence -ohjelmisto pitää sisällään vähintään ohjelmiston oman sisäänrakennetun käyttäjienhallintajärjestelmän, jossa käyttäjiä pystytään ryhmittelemään ja jakamaan eri käyttäjille tai ryhmille käyttöoikeuksia. Business intelligence -ohjelmistot voivat myös mahdollistaa jo olemassa olevien hakemistopalvelujen liittämisen ohjelmiston käyttäjähallintaan. Yleisimmin tuettuja ovat ainakin Windows Active Directory ja LDAP -hakemistopalvelut. Ulkoisia hakemistopalveluja käyttämällä pystytään organisaation eri ohjelmistojen käyttäjienhallinta toteuttamaan keskitetysti, eivätkä käyttäjät tarvitse useita eri tunnuksia ja salasanoja kirjaututtaessa eri järjestelmiin. (Peck 2008, 712–713.)

3.6 Muut perustyökalut

3.6.1 Mittaristot

Mittaristo-työkaluja (dashboards) käytetään organisaation määrittelemien tunnuslukujen seurantaan. Mittaristoilla pyritään hyvin visuaalisesti esittämään käyttäjälle hänen tarvitsemansa tiedot. Mittaristoissa käytetään erilaisia nuolia, trendiviivoja, palkkeja, karttoja ja värikoodeja osoittamaan liiketoiminnan eri toimintojen tilaa. Punainen väri mittaristossa kertoo tavallisesti negatiivisesta tilanteesta ja vastaavista vihreä positiivisesta. (Howson 2007, 45.)

Tyypillisesti mittaristoissa esitettävä tieto haetaan useista eri lähteistä. Ihanne-tapauksessa loppukäyttäjät pystyvät itse luomaan tarvitsemiaan mittareita, mutta kaikissa työkaluissa tämä ei ole mahdollista. Kehittyneimmissä mittaristo-työkaluissa mittaristot voivat päivittyä reaaliaikaisesti. (Howson 2007, 46.)

3.6.2 OLAP

OLAP (On-line Analytical Processing) eli moniulotteinen analysointi on business intelligence -ratkaisujen keskeisimpiä tiedon analysointiin tarkoitettuja työkaluja. OLAPia voidaan hyödyntää esimerkiksi budjetoinnissa, myynnin analysoinnissa ja ennustamisessa, markkinatutkimuksessa ja asiakasanalyysissä. OLAP-työkalujen kolme avainominaisuutta ovat: moniulotteiset näkymät tietoon, mahdollisuus monimutkaisiin laskutoimituksiin ja ajan älykäs hallinta. OLAPin moniulotteisuus kuvataan yleensä OLAP-kuutiona. (OLAP Council 1997.)

OLAP-kuution ulottuvuuksina voivat olla esimerkiksi tuoteryhmä, aika ja paikka. Kuution ulottuvuudet ovat tavallisesti hierarkisia, mikä mahdollistaa porautumisen ulottuvuuden tarkempiin tasoihin. Esimerkiksi tuoteryhmissä voidaan siirtyä yksittäisiin tuotteisiin ja ajassa kuukausiin, viikkoihin tai päiviin. (Hovi, Hervonen & Koistinen 2009, 93.) OLAPin tulee pystyä suoriutumaan yhtäaikaisista monimutkaisista laskutoimituksista tehokkaasti ja eri ulottuvuuksien läpi. Myyntiennusteita laskettaessa tulee voida laskea esimerkiksi liikkuvia keskiarvoja ja vuosikasvuja. OLAP-työkalussa tulee lisäksi pystyä määrittelemään erilaisia aikaan liittyviä käsitteitä ja suorittamaan aikaan liittyviä laskutoimituksia, kuten esimerkiksi laskea tase määritellyllä ajanjaksolla. (OLAP Council 1997.)

3.6.3 Tiedon louhinta

Tiedon louhinta (data mining) on OLAPin ohella toinen business intelligencen analysointityökalu. Tiedon louhinnalla pyritään suuresta datamäärästä löytämään piilossa olevaa odottamatonta ja arvokasta tietoa. Datamäärästä etsitään korrelaatioita, toistuvia ilmiöitä ja muita lainalaisuuksia. Tiedon louhinnassa voidaan käyttää useita erilaisia tekniikoita, kuten algoritmeja, koneoppimista, tilastotiedettä ja neuroverkkoja. Tiedon louhinnan tavoitteena

on, että löydettyä piilevää tietoa voitaisiin hyödyntää liiketoiminnan tukemiseen samaan tapaan kuin jo tiedossa ollutta tietoa. Tiedon louhintaa voidaan hyödyntää esimerkiksi asiakkaiden ostoskäyttäytymisen seuraamisessa, asiakasprofiilien luomisessa ja kysynnän ennustamisessa. (Hovi, Hervonen & Koistinen 2009, 98–99.)

4 AVOIN LÄHDEKOODI

Luvussa käydään lyhyesti läpi avoimen lähdekoodin historia ja synty. Lisäksi selvitetään avoimen lähdekoodin määritelmä sekä esitellään tutkittavissa sovelluksissa käytettävät avoimen lähdekoodin lisenssimallit ja niiden erityispiirteet.

4.1 Historia

Avoimen lähdekoodin esiasteena voidaan pitää ohjelmoija Richard Stallmanin vuonna 1984 perustamaa GNU-projektia. Projektin tavoitteena oli tuottaa vapaita ohjelmistoja (free software) ja lopulta julkaista täysin vapaa ja riippumaton käyttöjärjestelmä. Stallman perusti Vapaan ohjelmiston säätiön (Free Software Foundation) vuonna 1985 tukemaan vapaan ohjelmiston - liikettä ja GNU-projektia. (Moody 2001, 34–40.)

Säätiön julkaisema vapaan ohjelmiston määritelmä (The Free Software Definition 2009):

- Vapaus käyttää ohjelmaa mihin tahansa tarkoitukseen.
- Vapaus tutkia miten ohjelma toimii ja muuttaa sitä omiin tarpeisiin sopivaksi. Pääsy lähdekoodiin on tämän edellytys.
- Vapaus levittää kopioita ohjelmasta niin että voit auttaa naapuriasi.
- Vapaus parantaa ohjelmaa ja julkaista parannukset (ja paranneltu versio kokonaisuudessaan) yleisölle niin, että koko yhteisö hyötyy. Pääsy lähdekoodiin on tämän edellytys.

Netscape Communicationsin ilmoitettua julkistavansa suljetun www-selaimensa lähdekoodin kokoontui ryhmä arvostettuja hakkereita strategiapalaveriin helmikuussa 1998. Ryhmä arvioi, että Netscapen

ilmoituksen myötä yritysmaailma olisi valmiina kuulemaan hakkeriyhteisön ajatuksia ohjelmistojen avoimesta kehityksestä. Yritysmaailman kiinnostuksesta haluttiin saada kaikki irti. Palaveriin osallistujat päättivät myydä vapaista ohjelmistoista tuttua ideaa liiketoiminnan pohjalta. (Open Source Initiative 2009.) "Free software" (vapaa ohjelmisto) -termi päätettiin hylätä sen epäselvyyden takia. Englanninkielinen sana "free" tarkoittaa sekä vapautta että ilmaista. Ilmaisuudesta puhuminen ei kuulostanut hyvältä, kun tarkoituksena oli saada vakuutettua yritysmaailma. (Ingo 2005, 55.) Uudeksi käsitteeksi valikoitui selkeämpi avoin lähdekoodi (open source), ja samalla perustettiin voittoa tavoittelematon avoimen lähdekoodin etujärjestö, Open Source Initiative. Järjestön tarkoituksena on markkinoida ja hallinnoida uutta avoimen lähdekoodin määritelmää. (Open Source Initiative 2009.)

4.2 Määritelmä

Open Source Initiativen julkaisema Open Source Definition listaa kymmenen kohtaa, jotka lähdekoodin lisenssin on täytettävä, jotta se voidaan määritellä avoimeksi lähdekoodiksi. Kymmenes kohta lisättiin jälkikäteen vuonna 2004, mutta muutoin ehdot ovat pysyneet samana järjestön perustamisesta alkaen. (Open Source Initiative 2009.)

Open Source Initiativen avoimen lähdekoodin määritelmä (The Open Source Definition 2009):

1. **Vapaa levitysoikeus:** lisenssi ei saa rajoittaa ketään myymästä tai antamasta ohjelmaa osana ohjelmistoa, joka on koottu useasta eri lähteestä. Lisenssillä ei voida vaatia rojaltia tai maksua tämän kaltaisesta myynnistä.
2. **Lähdekoodi:** ohjelman tulee sisältää sen lähdekoodi, ja ohjelman levitys sekä lähdekoodina että käännettyssä muodossa tulee sallia. Ohjelman

levitessä ilman lähdekoodia tulee lähdekoodin olla helposti saatavilla kohtuullisin korvauksin, mielellään internetin kautta ilmaiseksi.

3. **Johdetut teokset:** lisenssin tulee sallia muokattujen ja johdettujen teosten luominen, ja niiden levittäminen on sallittava samoin lisenssiehdoin.
4. **Lähdekoodin yhteenkuuluvuus:** lisenssillä voidaan rajoittaa muokatun lähdekoodin levittämistä vain, jos lisenssi sallii korjaustiedostojen ja niiden lähdekoodin levittämisen. Lisenssillä voidaan vaatia, että johdetut teokset levitetään eri nimillä ja versionumeroilla kuin alkuperäiset teokset.
5. **Henkilöiden ja ryhmien syrjinnän kieltö:** lisenssillä ei voida syrjiä mitään henkilöä tai henkilöryhmää.
6. **Toimialojen syrjinnän kieltö:** lisenssillä ei voida kieltää ketään käyttämästä ohjelmaa tietyllä toimialalla.
7. **Lisenssin levittäminen:** ohjelmaan kuuluvien oikeuksien tulee suoraan soveltua kaikille, ilman jonkin uuden lisenssin käyttöönottoa.
8. **Lisenssi ei saa olla tuotekohtainen:** ohjelmaan kuuluvat oikeudet eivät saa riippua siitä, että ohjelma on osana tiettyä ohjelmistopakettia. Mikäli ohjelma erotetaan ohjelmistopaketistaan ja sitä käytetään ja levitetään ohjelmiston lisenssillä, tulee erotettavan ohjelman käyttäjien saada samat oikeudet kuin alkuperäisessä ohjelmistopaketissa.
9. **Lisenssi ei saa rajoittaa muita ohjelmia:** Lisenssi ei saa asettaa rajoituksia ohjelmille, joita levitetään lisensoidun ohjelman mukana.
10. **Lisenssin tulee olla teknisesti neutraali:** Lisenssin ehtoja ei voida muuttaa tietyn teknologian tai käyttöliittymän mukaan.

4.3 Lisenssit

Avoimen lähdekoodin ohjelmistojen levitys ja käyttö perustuu lisenssisopimukseen. Kaikissa avoimen lähdekoodin lisensseissä tekijänoikeus säilyy alkuperäisellä tekijällä. Lisenssisopimuksella tekijänoikeuden haltija antaa ainoastaan rajoitetun käyttöoikeuden ohjelmaansa. Eri lisenssisopimusten lisenssinsaaajille antamat oikeudet vaihtelevat hyvinkin paljon. Open Source Initiative ylläpitää kotisivuillaan listaa avoimen lähdekoodin määritelmän täyttäneistä ja järjestön hyväksymistä avoimen lähdekoodin lisensseistä. Lista käsittää tällä hetkellä vajaat 70 lisenssiä. (Open Source Initiative 2009.)

4.3.1 GNU General Public License

GNU General Public License (GNU GPL tai GPL) on Richard Stallmanin alunperin GNU-projektia varten kirjoittama lisenssi. Lisenssillä Stallman halusi varmistaa, että projektissa syntyneet ohjelmat säilyivät avoimina myös siirtyessään käyttäjältä toiselle. (Moody 2001, 43.) GPL-lisenssi on ylivoimaisesti suosituin avoimen lähdekoodin lisenssi. Verkon suurimman avoimen lähdekoodin ohjelmistoarkiston ja kehitysyhteisön SourceForgen projekteista noin puolet on GPL-lisenssoituja. (SourceForge 2010.)

GPL-lisenssin huomionarvoisin ehto liittyy lisenssin tarttuvuuteen. Lisenssi edellyttää, että jokainen teos, joka kokonaan tai osittain perustuu tai sisältää osia alkuperäisestä ohjelmasta tulee lisensoida kokonaisuudessaan ilman maksua samalla lisenssillä. Vaikka GPL-lisensoidusta ohjelmasta hyödynnettäisiin kuinka pientä osaa tahansa, tulee johdannainen lisensoida aina GPL-lisenssillä. Mikäli GPL-lisensoitu ohjelmaa yhdistetään toisen lisenssin ohjelman kanssa, korvaa GPL-lisenssi siihen yhdistetyn ohjelman lisenssin. GPL-lisenssi edellyttää myös, että vastaanottajalle ei voida asettaa mitään lisärajoituksia lisenssissä annettujen oikeuksien käytöstä. Tästä syystä

monet avoimen lähdekoodin lisenssit eivät ole yhteensopivia GPL-lisenssin kanssa. (Laurent 2004, 39–43.)

4.3.2 GNU Affero General Public License

GPL-lisenssiin perustuva GNU Affero General Public License (AGPL) on lähinnä verkkosovelluksia silmällä pitäen toteutettu lisenssi. AGPL-lisenssin keskeisin piirre on mahdollistaa sovelluksen lähdekoodin saaminen käyttöön, jos käyttäjä käyttää AGPL-lisenssoitua ohjelmaa verkon yli. (GNU Affero General Public License v3 2009.)

4.3.3 GNU Lesser General Public License

GNU Lesser General Public License (LGPL, tunnettiin aiemmin nimellä GNU Library General Public License) on etenkin ohjelmakirjastoissa käytetty pysyvä lisenssi, joka muistuttaa GPL-lisenssiä. Aivan kuten GPL-lisenssissä myös LGPL-lisensoiduista ohjelmista johdetut ohjelmat tulee lisensoida samalla lisenssillä. LGPL-lisenssi mahdollistaa kuitenkin LGPL-lisensoidun ohjelmakirjaston ja toisen ohjelman yhdistelmän lisensoinnin millä tahansa lisenssillä. Yhdistetyssä teoksessa ja kaikissa sen kopioissa tulee olla maininta hyödynnetyistä LGPL-lisensoiduista kirjastosta ja siitä, että kirjaston käytössä on noudatettava LGPL-lisenssin ehtoja. Lisäksi käyttäjällä on oltava mahdollisuus saada LGPL-kirjaston lähdekoodi haltuunsa ja muokata sitä haluamukseen, myös siinä tapauksessa, että kirjastoa olisi hyödynnetty osana suljettua ohjelmistoa. (Laurent 2004, 50, 55–57.)

4.3.4 Mozilla Public License

Netscapen julkistettua suljetun Communicator-www-selaimensa kooditiedostot kaikkien saataville vuonna 1998, kirjoitti se samalla kaksi uutta avoimen lähdekoodin lisenssiä vastaamaan yhtiön silloisia avoimen lähdekoodin suunnitelmia. Kirjoitetuista lisensseistä Netscape Public License

(NPL) antoi Netscape-yhtiölle ylimääräisiä oikeuksia muun muassa muuttaa kolmannen osapuolen kehittämän koodin lisenssiä. (Laurent 2004, 62–63.) Lisensseistä toista Mozilla Public Licenseä (MPL) pidetään erittäin hyvin kirjoitettuna, ja monet yhtiöt kuten Apple, Nokia ja Sun ovatkin käyttäneet MPL-lisenssiä pohjana omiin avoimen lähdekoodin lisensseihinsä (Laurent 2004, 84). MPL-lisenssin kehitystyö on kuulunut vuodesta 2003 asti Mozillaa säätiölle.

MPL on pysyvä lisenssi aivan kuten LGPL eli kaikki siitä tehtävät suorat johdannaiset on lisensoitava kokonaisuudessaan samalla MPL-lisenssillä. MPL mahdollistaa laajemman teoksen lisensoinnin millä tahansa lisenssillä. Laajemminkin teoksessa MPL-lisensoitu osa säilyttää oman lisenssinsä ja sen lähdekoodi on oltava saatavilla. On aina syytä tarkistaa, tulkitaanko uusi teos alkuperäisen johdannaiseksi vai omaksi laajemmaksi teokseksi. Laajemman teoksen MPL-osaan lisenssin haltija saa tekijänoikeuden alaiset oikeudet, jotka eivät sisällä patenti- ja tuotemerkko-oikeuksia. MPL-lisensoidun ohjelman patenttioikeudet ovat voimassa vain ohjelman muokkaamattomissa versioissa. (Laurent 2004, 67–73.)

5 TUTKIMUKSEN TOTEUTUS

Tutkimukseen tarvittiin kaksi avoimen lähdekoodin business intelligence -ratkaisua, jotka vastaavat ominaisuuksiltaan kaupallisia vaihtoehtoja mahdollisimman hyvin ja mahdollistavat sekä raporttien luomisen että niiden jakelun. Tutkimukseen valittavia eri vaihtoehtoja pyrittiin selvittämään avoimen lähdekoodin projektien omien kotisivujen, olemassa olevien tutkimusten sekä alan uutisoinnin ja artikkeleiden perusteella. Selvityksen perusteella havaittiin hyvin nopeasti, että potentiaalisia avoimen lähdekoodin kokonaisuuksia, joista voisi mahdollisesti olla korvaamaan kaupalliset tuotteet, on tällä hetkellä kaksi. Pelkästään raporttien luomiseen tarkoitettuja avoimen lähdekoodin ratkaisuja on olemassa muutamia, mutta nämä eivät mahdollistaneet raporttien jakelua ja näin ollen eivät voineet tulla valituiksi. Avoimen lähdekoodin business intelligence -projekteja on ollut useampiakin, mutta niiden kehitys on loppunut ja viimeisimmät julkaisut tai päivitykset ovat useamman vuoden takaa.

Tutkimukseen valitut avoimen lähdekoodin business intelligence -ratkaisut versiotietoineen:

- JasperReports 3.7.0 -luokkakirjasto ja iReport 3.7.0 -käyttöliittymä
- JasperServer 3.7.0 -raportointipalvelin
- Pentaho Report Designer 3.6.0 -raportointiohjelmisto
- Pentaho Business Intelligence Server 3.5.0 -raportointipalvelin

JasperReports ja Pentaho ovat ensisilmäyksellä varsin samankaltaisia ohjelmistoja. Molemmat ovat hyvin kattavia business intelligence -

ohjelmistokokonaisuuksia jotka raportoinnin lisäksi mahdollistavat myös tiedon integroinnin ja analysoinnin. Tällä hetkellä molempien kehitystyö jatkuu aktiivisesti ja viimeisimmät julkaisut ovat alkuvuodelta 2010.

Kumpaisestakin ratkaisusta ovat tarjolla myös kaupalliset versiot, mutta tutkimukseen valittiin molemmista vapaasti saatavilla olevat ns. community-versiot.

Valitut business intelligence ratkaisut asennettiin tutkimusympäristönä toimineisiin Windows XP - ja Windows 7 -käyttöjärjestelmiin. Kumpaankin käyttöjärjestelmään asennettiin paikallisesti Apache Tomcat 5.5.20 -sovelluspalvelin sekä MySQL 5.1.36 -relaatiotietokanta. Molempien business intelligence -ratkaisujen mukana saatiin tietokantaan runsaasti mallidataa, jota voitiin hyödyntää ratkaisujen raportointiominaisuuksia tutkittaessa.

5.1 JasperReports

Vuonna 2001 yhden miehen projektina alkaneen JasperReportsin takana on Jaspersoft-niminen yhdysvaltalainen yhtiö. Yhtiö vastaa kaupallisen Jaspersoft Business Intelligence Suite -ohjelmiston kehityksestä.

Ohjelmistopakettiin kuuluvat JasperReportsin lisäksi JasperServer, JasperAnalysis, JasperETL ja iReport tuotteet. Kaikista yhtiön kaupallisista tuotteista on vapaasti saatavilla avoimen lähdekoodin versiot. Jaspersoft valvoo itse JasperForge-nimellä kulkevaa avoimen lähdekoodin yhteisöä. JasperForge-yhteisössä on 120 000 jäsentä, yli 350 projektia, ja projekteja on ladattu yhteensä yli 10 miljoonaa kertaa. Kaupallisia asiakkaita Jaspersoft on ilmoittanut olevan 12 000. JasperReports-kirjasto on lisensoitu LGPL-lisenssillä, iReport GPL-lisenssillä ja JasperServer AGPL-lisenssillä.

(JasperForge 2009.)

5.2 Pentaho

Pentaho on vuonna 2004 perustettu ohjelmistoyhtiö. Pentahon taustalta löytyy alan veteraaneja, jotka ovat ennen Pentahoa olleet tekemässä kaupallisia business intelligence -tuotteita muun muassa Business Objectsille, IBM:lle ja Oraclelle (About Pentaho 2009). Pentahon kaupallinen Pentaho Business Intelligence Suite Enterprise Edition sisältää Pentaho Reporting, Analysis, Dashboards, Data Integration ja Data Mining -tuotteet. Kaupallisista tuotteista on saatavilla vastaavat avoimen lähdekoodin versiot. Pentahon tuotteita on ladattu yli viisi miljoonaa kertaa ja yhtiö mainostaa sivullaan sillä olevan yli 8 000 yritysasiakasta (Pentaho 2009). Pentaho-tuotteet ovat suurimmaksi osaksi lisensoitu GPL-lisenssillä, mutta osa komponenteista on lisensoitu muun muassa LGPL- ja MPL-lisenssein (Pentaho - Business Intelligence 2010).

6 TUTKIMUKSEN TULOKSET

Luvussa käydään läpi tutkittavana olleet ohjelmat ja nostetaan esiin niiden oleellimmat ominaisuudet sekä mahdolliset puutteet ja esitetään havaintoja ohjelmien yleisestä toiminnasta. Luvussa selvitetään myös millaisia eroavaisuuksia tutkittavista tuotteista löytyy.

6.1 Raportointiohjelmat: JasperReports / iReport ja Pentaho Report Designer

6.1.1 Esittely

JasperReports on kokonaan Javalla toteutettu raporttien luomiseen tarkoitettu luokkakirjasto. JasperReportsissa ei itsessään ole lainkaan käyttöliittymää, vaan luokkakirjastoa voidaan tavallisten java-kirjastojen tapaan hyödyntää muissa java-sovelluksissa. JasperReportsille on vuosien aikana toteutettu muutamia eri käyttöliittymiä. iReport-nimellä kulkeva käyttöliittymä on tällä hetkellä ainoa, joka hyödyntää täysin JasperReports-kirjastoa. Uudet versiot JasperReportsista ja iReportista julkaistaan aina samanaikaisesti, ja niillä on tätä nykyä yhtenevät versionumeroinnit.

Pentaho Report Designer on raporttien suunnitteluun ja luomiseen käytettävä työpöytäohjelma. Pentaho Report Designer on niin ikään toteutettu kokonaan Java-kielellä. Tutkittavana olleen Report Designerin 3.6 -versiota edeltäneen 3.5 -version lähdekoodi kirjoitettiin kokonaan uudelleen.

Sekä JasperReportsissa että Pentahossa raportit määritellään XML-kielellä. JasperReports-kirjastoa käytettäessä voidaan raportit luoda ja määritellä periaatteessa millä tahansa tekstieditorilla. iReport-ohjelmassa käyttäjä voi valita, käyttääkö raporttien luomiseen iReportin graafista käyttöliittymää vai

ohjelmaan sisäänrakennettua tekstieditoria. Pentaho Report Designer ei sisällä teksti- tai XML-editoria, joten raporttien määrittelyä ei voida suorittaa lainkaan käsin, vaan raporttien suunnittelu ja määrittely on toteutettava Report Designerin käyttöliittymän avulla.

6.1.2 Asennus

iReportista on valmiit asennuspaketit Windowsille (XP, Vista ja 7), Macille (OS X) sekä Linuxille. iReport on myös mahdollista asentaa lisäosaksi NetBeans IDE -sovelluskehittimeen. iReportin asennus on hyvin yksinkertainen, eikä käyttäjän tarvitse tehdä muuta kuin valita haluamansa kohdehakemisto asennukselle. iReportia käytettäessä JasperReportsia ei tarvitse asentaa erikseen, vaan se sisältyy iReportin asennuspakettiin.

Pentaho Report Designerista on niin ikään ladattavissa yhteensä kolme tiedostopakettia, yksi kullekin käyttöjärjestelmälle (Windows XP tai uudempi, OS X ja Linux) sekä ohjelman pelkkä lähdekoodi. Tiedostopaketteihin kuuluu varsinaisen ohjelman ja lisenssiehtojen lisäksi muutama valmis raporttipohja sekä parikymmentä malliraporttia. Pentaho Report Designeria ei tarvitse asentaa ollenkaan, vaan ohjelma voidaan käynnistää suoraan ajamalla tiedostopaketista löytyvä jar-tiedosto. Sekä iReport että Pentaho Report Designer vaativat toimiakseen Java 1.5.0 -version tai sitä uudemman.

6.1.3 Datayhteydet

Ohjelmissa on hyvin kattavat ominaisuudet käyttää lukuisia eri tietolähteitä. Molemmista ohjelmista löytyy tuki useille kymmenille eri relaatiotietokannoille, mukaan lukien kaikki yleisimmin käytössä olevat tietokannat, kuten IBM DB2, Oracle, MySQL, PostgreSQL, MS Access ja MS SQL Server. Tietokantayhteyttä luotaessa käyttäjän tulee syöttää tietokantapalvelimen osoite, tietokannan nimi, käyttäjätunnus sekä salasana

(ks. KUVIO 2). iReportissa ja Pentahossa voidaan tietojen syöttämisen jälkeen vielä testata, että kaikki tiedot on syötetty oikein ja yhteys saadaan muodostettua. Relaatiotietokantojen lisäksi molemmilla ohjelmilla voidaan luoda datayhteys myös muun muassa XML-, Microsoft Excel- ja CSV-tiedostoihin sekä OLAP-tietokuutioihin.

KUVIO 2. Tietokantayhteyden luominen iReportissa

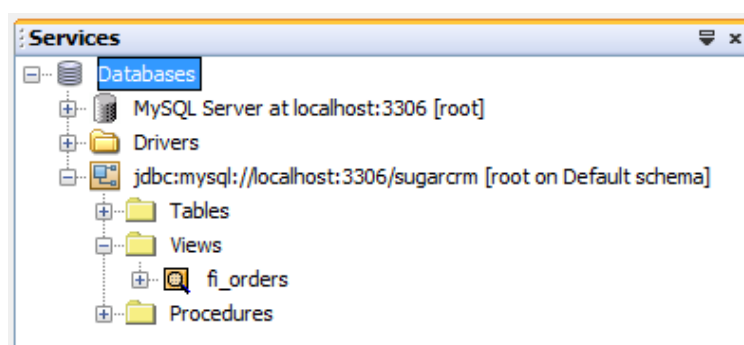
iReportissa on valittavana kaiken kaikkiaan noin 20 JDBC-ajuria, ja kaikki yhteydet relaatiotietokantoihin toteutetaan juuri JDBC-rajapinnan kautta. iReportissa voidaan myös käyttää JDBC-ODBC-siltaa, jolla saadaan yhteys ODBC-rajapinnan tarjoaviin tietolähteisiin. Tietolähteisiin, joita JasperReports ei suoraan tue, voidaan luoda yhteys erityisen Custom JRDataSource -rajapinnan avulla. Tämän rajapinnan kautta saadaan yhteys käytännössä ainutlaatuisiin ja varta vasten yritystä tai organisaatiota varten luotuihin

tietolähteisiin, mutta tällöin kehittäjän tulee kirjoittaa yhteyden muodostamisessa tarvittava ohjelmakoodi itse. JasperReports mahdollistaa XML-tiedostojen käytön tietolähteenä paitsi paikallisesti myös verkon ylitse.

Pentaho Report Designerissa on valittavana datayhteys lähes 40 relaatiotietokantaan. Aivan muutamaa poikkeusta lukuun ottamatta käyttäjä voi valita, luodaanko tietokantayhteys joko natiivin JDBC-rajapinnan kautta tai ODBC-rajapintaa hyödyntämällä. Yhtenä poikkeuksena tähän on MS Access, johon saadaan yhteys ainostaan ODBC-rajapinnalla.

6.1.4 Tietolähteiden käsittely

Raporttien tietolähteenä käytettäviin relaatiotietokantoihin päästään käsiksi sekä iReportissa että Pentahossa tavallisilla SQL-kyselyillä ja proseduureilla. Kuviossa 3 nähdään, kuinka iReportin Services-paneeli esittää tietokannan taulut, näkymät ja proseduurit eriteltyinä selkeästi omissa hakemistoissaan. Paneelin kautta pystytään luomaan uusia tauluja, näkymiä ja proseduureja sekä poistamaan olemassa olevia. Kaikki lisäykset ja poistot tietokantaan voidaan kirjoittaa käsin SQL-kyselyinä.



KUVIO 3. iReportin Services-paneeli

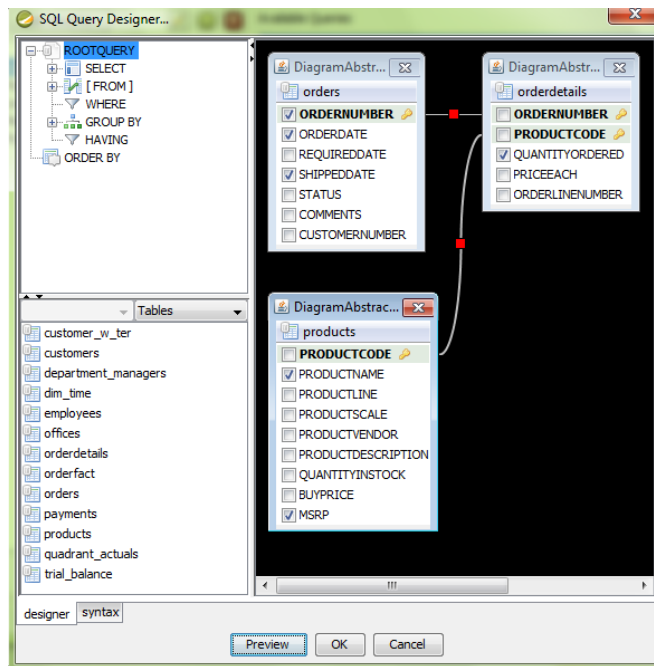
Paneelin kautta voidaan nähdä suoraan myös taulujen, näkymien tai yksittäisten sarakkeiden sisältämät tiedot parilla hiiren klikkauksella, ilman että minkäänlaisia kyselylauseita tarvitsee kirjoittaa. Varsinaiset tietokannan

hakukyselyt kirjoitetaan iReportin Report query -ikkunassa. Kyselykielistä JasperReports tukee natiivisti seuraavia: SQL, HQL, PL/SQL, EJBQL, MDX, XMLA-MDX, XPath ja XPath 2.0. Hakukyselyt voidaan myös tallentaa ja vastaavasti jo tallennetut kyselyt voidaan ladata käyttöön. Ikkunassa nähdään automaattisesti haun palauttamien sarakkeet ja niiden tietotyypit.

Pentaho Report Designer ei mahdollista tietokannan muokkaamista. Report Designerin avulla ei kyetä poistamaan tai luomaan uusia tauluja, näkymiä tai prosedureja. Näkymiä ja prosedureja voidaan hyödyntää hakukyselyissä, mutta Report Designer ei tarjoa minkäänlaista hakemistopuuta tai listausta, jossa tietokannan näkymät ja proseduurit olisivat nähtävissä. Hakukyselyt Pentahossa voidaan suorittaa SQL, MQL, XQuery ja MDX -kielillä.

Hakukyselyn kirjoittamisen yhteydessä kyselyn palauttamien tietojen voidaan esikatsella. Sekä iReportissa että Pentaho Report Designerissa hakukyselyiden luomisen apuna voidaan myös käyttää graafista työkalua (ks. kuvio 4).

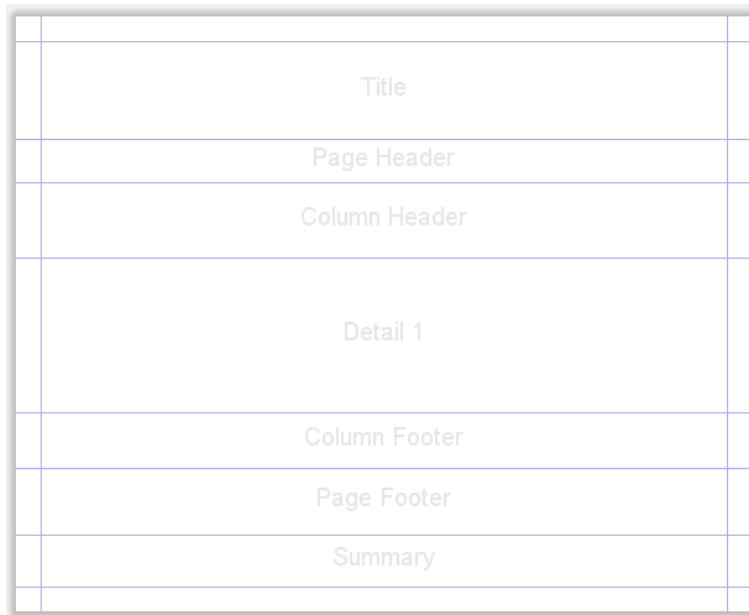
Kyselyn suunnittelutyökalu on iReportissa ja Pentahossa lähes täysin identtinen sekä toiminnoiltaan että ulkoasultaan. Ainoana erona on Pentahon mahdollisuus kyselyn esikatseluun.



KUVIO 4. Query Designer on graafinen työkalu hakukyselyiden luomiseen.

6.1.5 Raporttien rakenne

Sekä JasperReportsin että Pentahon raportit koostuvat muutamasti erilaisesta horisontaalisista osiosta (band) (ks. kuvio 5). Kaikki osiot ovat koko raportin levyisiä, mutta niiden korkeutta voidaan muuttaa ja tarpeettomaksi katsotut osiot voidaan joko poistaa tai piilottaa kokonaan. Valmista raporttia katseltaessa tai raporttia tulostettaessa nämä raportin eri osiot käyttäytyvät hieman eri tavalla.



KUVIO 5. Raportin perusrakenne JasperReportsissa

JasperReportsissa eri osioita on kaiken kaikkiaan 11 ja Pentahossa yhdeksän. Kummassakin ohjelmassa ovat omat ylä- ja alatunniste-osiot raportin sivujen lisäksi myös tietosarakkeille sekä ryhmittelyille. Group header ja -footer osioita ei ole käytössä lainkaan, jos raportissa ei käytetä tiedon ryhmittelyä, ja vastaavasti jos tietoa ryhmitellään, niin ryhmittelyosiot luodaan jokaiselle ryhmittelytasolle erikseen. Detail-osio on raportin varsinainen pääosio. Suurin osa raportin kaikista tulostuselementeistä sijoitetaan juuri Group ja Detail -osioihin. Detail-osio jatkuu raportin sivulta toiselle, jos tulostettavia tietoja on niin runsaasti, että raportti ei mahdu yhdelle sivulle. JasperReportsin background- ja Pentahon watermark-osioihin voidaan lisätä esimerkiksi kuvia ja kuvioiteja, joiden halutaan näkyvän raportin jokaisen sivun taustassa.

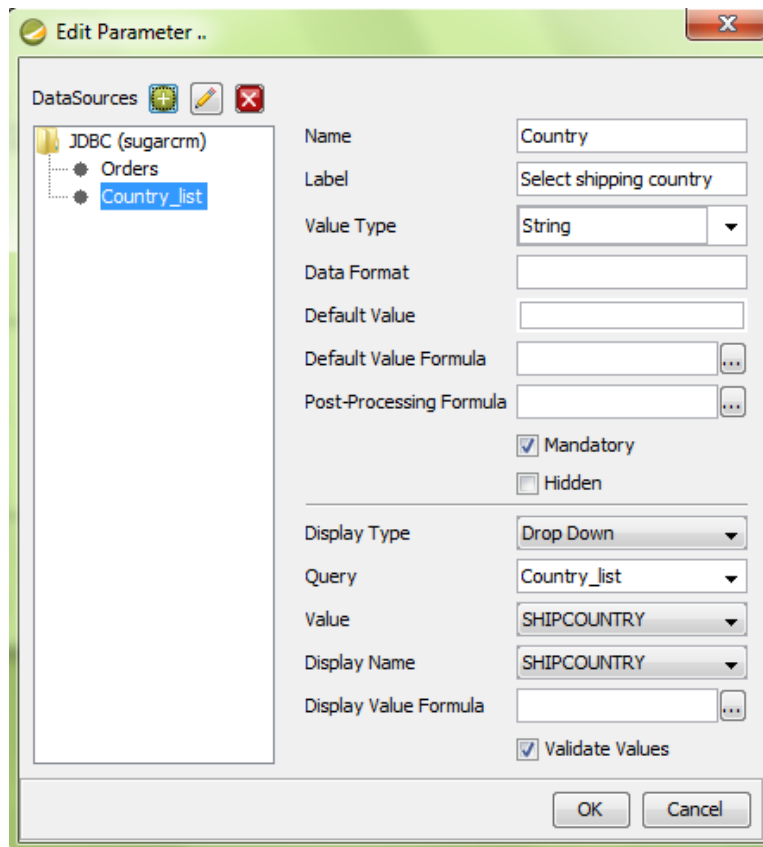
JasperReportsin kolme osiota, joita Pentahosta ei löydy, ovat title, last page footer ja summary. Title-osio ja sen sisältö näkyvät ainoastaan raportin alussa ja tämä osio voi olla myös kokonaan erillinen sivu raportissa. Last page footer on erityinen raportin viimeisen sivun alatunniste ja summary-osiota voidaan nimensä mukaisesti käyttää raportin yhteenvedona. Summary-osio näytetään

ainoastaan raportin lopussa. Pentahon osio, jota ei puolestaan JasperReportista löydy, on crosstab (ristiintaulukointi). Pentahossa crosstab-osio tulee näkeviin ainoastaan, jos raporttiin tehdään ristiintaulukoita.

6.1.6 Parametrit

Luotaessa uutta parametriä parametri tulee nimetä, sille täytyy asettaa luokka/tietotyyppi ja haluttaessa asettaa oletusarvo. Parametrin oletusarvo voidaan määrittää kaavan avulla. Jotta raportin loppukäyttäjä pystyy helposti syöttämään parametrille arvon tai arvoja, täytyy parametrille määrittää myös arvon syöttämisessä käytettävä elementti. Elementti voi olla esimerkiksi tavallinen tekstikenttä, alavetovalikko, monivalintakenttä, radiopainike, monivalintaruutu tai päivämäärän valitsin. Parametrille tulee myös määrittää sitä koskeva tietolähde sekä tarvittaessa syötetään parametrille hakukysely, jolla elementtiin saadaan dynaamisesti haettua tarvittavat tiedot. Parametrit voidaan haluttaessa asettaa pakollisiksi tai piilottaa käyttäjältä.

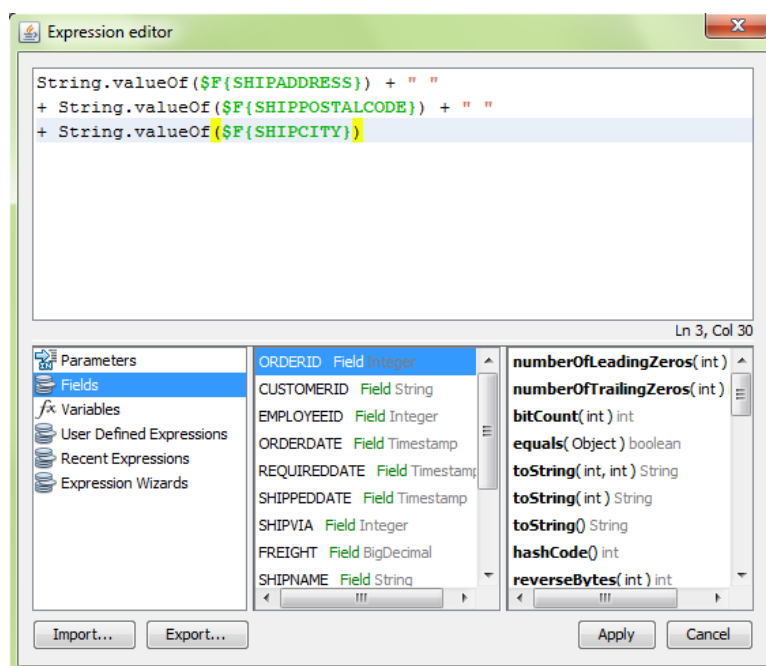
Parametrien määrittäminen eroaa iReportissa ja Pentahossa siten, että iReportissa käyttäjä voi ainoastaan nimetä parametrin, asettaa luokan ja oletusarvon. Käytännössä kaikki muut määrytykset parametreihin tehdään JasperServerin puolella. Pentaho Report Designerissa parametrien määrittäminen hoidetaan samassa ikkunassa (ks. kuvio 6).



KUVIO 6. Parametrin määrittäminen Pentaho Report Designerissa

6.1.7 Kaavat

Sekä JasperReportsin että Pentahon raporteissa voidaan käyttää kaavoja (expression / formula). Kaavoissa voidaan viitata raportin kenttiin, muuttujiin ja parametreihin. Kaavoilla voidaan esimerkiksi suorittaa laskutoimituksia, luoda uutta sisältöä tietolähteistä saatujen tietojen pohjalta, yhdistellä tietoja sekä muuttaa eri objektien ulkoasumäärittäyksiä. Esimerkiksi osoitetiedot eritellään usein katuosoitteeksi, postinumeroksi ja -toimipaikaksi tallennettaessa tietoja tietokantaan. Kaavalla nämä osoitteen eri kentät voidaan yhdistää yhdeksi kentäksi (ks. kuvio 7). iReportista ja Pentaho Report Designerista löytyy kummastakin editori kaavojen määrittämiseen.



KUVIO 7. iReportin kaavaeditori

6.1.8 Tiedon ryhmittely ja lajittelu

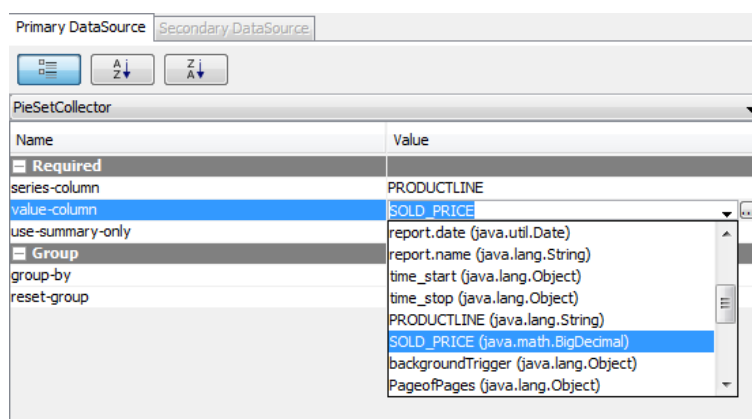
iReportissa ja Pentahossa on mahdollista luoda ryhmiä käytännössä rajattomasti. Ryhmittelyt voidaan tehdä kenttien, muuttujien, parametrien tai erillisen kaavan perusteella. Jokaiselle luodulle ryhmälle luodaan oma header- ja footer -osio. Jotta ryhmittely toimisi halutulla tavalla, täytyy tiedot lajitella oikein jo siinä vaiheessa kun tiedot haetaan tietolähteistä. Lajittelu voidaan toteuttaa joko nousevasti tai laskevasti, ja lajittelutasoja voidaan luoda rajattomasti. Tiedon hakuihin voidaan myös lisätä erilaisia rajauksia kaavoja käyttämällä. Rajauksissa käytettävien kaavojen tulee palauttaa joko true tai false toimiakseen.

6.1.9 Kaaviot

JasperReportsista löytyy parikymmentä eri kaaviomallia. Valittavana on esimerkiksi sektorikaavio, muutama erityyppinen pylväskaavio, viivakuvaajia, kuplakuvat, hajontakuviot, aluekaavio, Ganttin kaavio ja mittari. Kaavioiden lähdetietoihin voidaan tavallisten tietokenttien lisäksi

lisätä myös muuttujia, parametrejä ja kaavoja. Kaavioiden ulkoasu ja ominaisuudet ovat hyvin muokattavissa.

Pentaho Report Designerissa on hieman JasperReportsia vähemmän kaaviomalleja. Kaavioita on silti toistakymmentä ja kaikki yleisimmät tyypit ovat Pentahossakin käytettävissä. Erilaisten pylväs-, sektori- ja aluekaavioiden sekä viivakuvaajien lisäksi käyttöön voidaan ottaa kuplakuvia, hajontakuviot, vesiputous- tai tutkakaavio. Kuvio 8 näyttää kuinka kaavion lähdetiedot määritellään Pentahossa. Myös kaavioiden ominaisuudet määritellään samassa ikkunassa.



KUVIO 8. Kaavion lähdetietojen määrittäminen Pentahossa

6.1.10 Ristiintaulukointi

JasperReportsin ristiintaulukointi on toteutettu varsin selkeästi iReportin käyttöliittymässä. Crosstab-elementti tulee sijoittaa raportin summary-osioon. Käyttäjän pudottaessa elementin raporttiin avautuu "new crosstab" -ikkuna. Aivan ensimmäisenä käyttäjän tulee valita mitä tietojoukkoa ristiintaulukoinnissa halutaan käyttää. Tietojoukon valinnan jälkeen määritellään taulukon riveille ja sarakkeille halutut muuttujat. Muuttujia voidaan ryhmitellä muuttujan tietotyypistä riippuen. Esimerkiksi timestamp-tyyppiset muuttujat voidaan ryhmitellä vuoden, kuukauden tai vaikkapa päivän mukaan. Seuraavaksi määritetään mitä ristiintaulukossa lasketaan, eli

valitaan muuttuja taulukon datakenttään. Oletuksena datakentän arvojen laskentaperusteena on "count", lukumäärän laskeminen. Valittavina ovat kuitenkin kaikki perinteiset laskentaperusteet kuten summaus, keskiarvo, varianssi, suurin ja pienin. Lopuksi voidaan taulukkoon asettaa vielä jokin lukuisista valmiista väriteemoista ja samalla on mahdollisuus lisätä taulukkoon automaattisesti vielä rivi ja sarake kokonaismäärien laskemiseen.

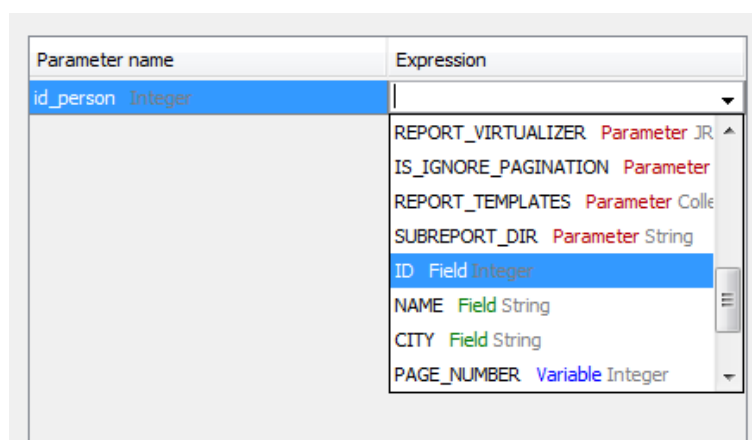
Kun ohjattu ristiintaulukointi on saatu vietyä loppuun, ilmestyy iReporttiin oma välilehti juuri luodulle ristiintaulukolle, jossa taulukkoon on helppo tehdä muutoksia ja lisäyksiä. Varsinaisessa raportin pääsuunnittelutilassa ristiintaulukointi näkyy ainoastaan staattisena kuvakkeena. Ohjatulla toiminnolla voidaan helposti luoda varsin yksinkertainen ristiintaulukko, jota käyttäjä pystyy myöhemmin manuaalisesti muokkaamaan. Ohjatussa näkymässä käyttäjä voi esimerkiksi määrittää taulukon rivi- ja sarakemuuttujille vain kaksi ryhmittelytasoa, mutta käsin tasoja voidaan lisätä useita. Ohjatussa näkymässä taulukon otsikoihin ei voida puuttua, joten ne voi olla syytä muokata jälkeinpäin ja mahdollisesti lisätä otsikkotasoja. Muuttujien avulla taulukkoon voidaan lisätä myös laskentakaavoja, kuten esimerkiksi taulukon lukumäärien oheen prosenttiosuudet.

Pentaho Report Designer 3.6:ssa ristiintaulukointi on vasta kokeiluasteella ja ominaisuus on vielä hyvin keskeneräinen. Ristiintaulukointiominaisuus ei ole ohjelmassa oletuksena edes käytössä, vaan käyttäjän tulee erikseen ottaa se käyttöön ohjelman ominaisuusvalikosta. Kun ristiintaulukointi on otettu käyttöön, tulee käyttäjän lisätä raporttiin ristiintaulukkoryhmä, jolloin raporttiin lisätään myös crosstab-osio. Ristiintaulukon määrittelemisen tapahtuu kokonaan käsin Pentaho Report Designerin structure-paneelissa. Käyttäjän tulee määrittää ryhmittelyt ja otsikot taulukon rivi- ja sarakemuuttujille sekä taulukon tietokenttä. Yksinkertaisinkaan mahdollinen ristiintaulukko ei tulostu suoraan oikein, vaan käyttäjä joutuu lisäämään

yhden tyhjän elementin taulukkoon, jotta sarakkeet menevät kohdalleen. Rivien ja sarakkeiden summaukset ja muut laskutoimitukset eivät ainakaan vielä ristiintaulukoinnin kokeiluversiossa onnistu.

6.1.11 Aliraportit

iReportissa aliraportti lisätään raporttiin raahaamalla subreport-elementti raportin päälle, jolloin aliraportti-ikkuna avautuu. Avautuvassa ikkunassa käyttäjä voi valita, rakentaako hän kokonaan uuden raportin aliraportiksi vai käyttääkö jo olemassa olevaa tallennettua raporttia. Raporttiin voidaan niin haluttaessa lisätä myös pelkästään tyhjä aliraportti-elementti, jolloin käyttäjä voi määritellä kaiken käsin. Aliraportit ovat JasperReportsissa itsenäisiä raportteja eli niitä on mahdollista lukea ja käsitellä myös ilman pääraporttia. Aliraportteja voidaan lisätä raportteihin useita ja aliraportteihin voidaan valita pääraportista poikkeavat datayhteydet. Pääraportista poikkeavat datayhteydet tulee määritellä kaavaan. Pääraportin sisältö voidaan hakea esimerkiksi relaatiotietokannasta ja aliraportti muodostaa yksittäisestä xml-dokumentista. JasperReports mahdollistaa parametrien liittämisen aliraportteihin esimerkiksi niin, että aliraportin parametrit haetaan pääraportin perusteella (ks. kuvio 9).



KUVIO 9. Parametrin liittämisen aliraporttiin iReportissa

Myös Pentahossa aliraportti lisätään raahaamalla subreport-elementti

raporttiin. Ohjelma kysyy käyttäjältä, luodaanko "inline" vai "banded" - aliraportti. Käytettäessä inline-vaihtoehtoa aliraportti vie raportista käyttäjän ennalta määrittelemän tilan, jolloin pystytään helposti asettelemaan useita aliraportteja esimerkiksi vierekkäin. "Banded"-valinnalla aliraportin koko muuttuu dynaamisesti sen mukaan, kuinka paljon tietoa aliraportin kysely palauttaa. Pentaho ei anna käyttäjälle mahdollisuutta käyttää aliraporttina jo olemassa olevaa raporttia vaan aliraportti on luotava sitä lisättäessä.

Pentahossa aliraportti kuuluu aina osaksi pääraporttia, eikä aliraporttia voida tallentaa tai käsitellä omana raporttinaan. Pentahon aliraportteihin voidaan liittää parametrejä ja pääraportista poikkeava datayhteys.

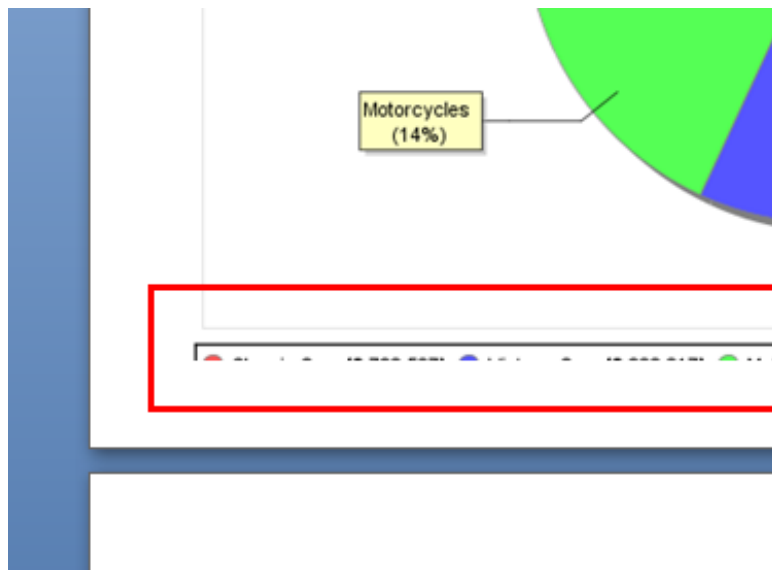
6.1.12 Raporttien tallennus

JasperReportsissa kaikki raporttien määrytykset tallennetaan sen omassa jrxml-tiedostoformaattissa. Pentaho Report Designerin raporttimäärytykset tallennetaan puolestaan prpt-tiedostoformaattissa. Molempien sovellusten raportteja voidaan kääntää kaikkiin yleisimpiin tiedostoformaatteihin. Tuettuina ovat pdf, rtf, html, xls ja csv -tiedostoformaattit. Näiden lisäksi JasperReports tukee vielä odt ja xml -formaatteja. Sekä iReportissa että Pentahossa raportit voidaan esikatsella kaikissa ohjelmien tukemissa tiedostoformaateissa.

iReportissa raporttien kääntäminen eri tiedostoformaatteihin tapahtuu juuri ohjelman esikatselutilassa. Kaikki raportit voidaan esikatsella joko ulkoisissa ohjelmissa, kuten Microsoft Excelissä tai internetselaimessa, tai suoraan iReportin sisällä. Kaikista parhaiten raportit kääntyvät pdf-formaattiin. Tällöin kaikki raporttiin suunnitellut muotoilut tulostuvat halutulla tavalla. Raportit kääntyvät varsin hyvin myös muihin formaatteihin, mutta esimerkiksi xls-käännöksissä on havaittavissa, että excel-taulukon solut eivät aina tulostu oikean kokoisiksi. Kaikki tieto menee kyllä oikein läpi, mutta esimerkiksi osa teksteistä saattaa olla vain osittain näkyvissä tai rivitetty

automaattisesti. Html-käännöksissä kaikki raporttien muotoilut tallennetaan suoraan html-koodin sisään.

Pentaho Report Designerissa raporttien kääntäminen hoidetaan suoraan file-valikosta. Raporttien muotoilut siirtyvät oikein hyvin xls-formaattiin, mutta parhaiten myös Pentahon käännöksistä raportit tulostuvat pdf-formaatissa. Rtf-käännöksissä muotoilut eivät täysin toimi, vaan raporttiin saattaa ilmestyä turhia tyhjiä rivejä. Tällöin raportin sivutukset eivät välttämättä tulostu suunnitellulla tavalla, vaikka kaikki raportin asetukset olisikin määritetty oikein. Sivutusongelmat aiheuttavat myös sen, että raportin kuva-objektit saattavat leikkaantua reunoista kuten kuviosta 10 voidaan nähdä. Html-käännöksessä Pentaho luo raportille erillisen tyylitiedoston, mikä helpottaa raportin muokkaamista jälkikäteen huomattavasti. Erillisestä tyylitiedostosta huolimatta aivan kaikkia raportin tyylimäärittelyksiä ei kuitenkaan tallenneta tyylitiedostoon, vaan osa määrittelyistä on html-koodin seassa.



KUVIO 10. Pentahon raporttia käännettäessä kaavion selite leikkaantuu.

6.2 Jakelujärjestelmät: JasperServer ja Pentaho Business Intelligence Server

JasperServer ja Pentaho Business Intelligence Server ovat raporttien jakelujärjestelmiä, joilla jaetaan ja välitetään raportointiohjelmistoilla luotuja raportteja raporttien loppukäyttäjille. Käyttäjät kirjautuvat jakelujärjestelmiin omilla tunnuksillaan ja pääsevät näin käsiksi niihin raportteihin, joihin omaavat käyttöoikeudet.

6.2.1 Asennus

JasperServerin perusasennukseen on tarjolla valmiit ajettavat asennustiedostot Windowsille (XP, Vista ja 7) sekä Linuxille. Asennustiedosto sisältää JasperServerin WAR-tiedoston, iReport -raportointiohjelmiston, Tomcat-sovelluspalvelinohjelmiston, MySQL-relaatiotietokantaohjelmiston sekä hieman mallidataa ja -raportteja. Perusasennuksessa voidaan valita mitkä ohjelmistot oletaan käyttöön. Jos käytöstä löytyy entuudestaan jo Tomcat-palvelin ja MySQL-tietokanta, voidaan JasperServer määritellä toimimaan niissä. Mikäli JasperServer halutaan asentaa muuhun kuin Tomcat+MySQL-ympäristöön, tulee asennus suorittaa manuaalisesti WAR-tiedostopakettin kautta. Manuaaliasennuksessa tuettuina ovat Tomcat-, JBoss- ja GlashFish -sovelluspalvelimet sekä MySQL- ja PostgreSQL -tietokannat.

Pentaho Business Intelligence Serveristä ei ole valmista ajettavaa asennuspakettia vaan asennus tulee suorittaa manuaalisesti. Pentahosta sekä Windowsille että Linuxille saatavat asennuspaketit on oletuksena konfiguroita HSQLDB -tietokannalle. Käyttäjän tulee muuttaa muutamaa Pentahon konfiguraatitiedostoa, jotta järjestelmä saadaan toimimaan toisella tietokannalla. Tehtävät muutokset koskevat lähinnä tietokannan käyttämään JDBC-ajuria. Perusasennuksessa muutosten tekeminen vie käyttäjältä vain

muutaman minuutin. Pentaho voidaan konfiguroida kaikille yleisimmille sovelluspalvelimille ja relaatiotietokannoille.

6.2.2 Raporttien asennus

Kaikkien JasperServeriin asennettavien raporttien tulee olla JasperReportsin omassa jrxml-tiedostoformaattissa. Raporttien asennus on jaettu seuraavaan kuuteen vaiheeseen: naming, jrxml, controls & resources, data source, query ja customization. Aivan ensimmäisenä raportti tulee nimetä ja samalla raporttiin voidaan liittää kuvaus, joka näkyy käyttäjille raporttilistan yhteydessä.

Toisessa vaiheessa haetaan haluttu raportti jrxml-tiedostona. Seuraavaksi raporttiin tulee lisätä kaikki tarvittavat kontrollielementit, joita tarvitaan raportin parametrien syöttämiseen. Samassa yhteydessä voidaan myös määritellä raportin muut resurssit, tyypillisesti muita resursseja ovat muun muassa raportissa esiintyvät logot ja kuvat. Tämän jälkeen tulee määritellä raportin käyttämä datayhteys. Raportin asennuksen yhteydessä voidaan myös määritellä raporttiin hakukysely. Asennettava raportti sisältää normaalisti jo hakukyselyn ja kysely on tallennettuna raportin jrxml-tiedostossa, joten niitä ei yleensä jouduta enää määrittelemään. Jos raportin hakukysely kuitenkin kirjoitetaan vielä asennuksen yhteydessä uudelleen, korvaa uusi kysely jo olemassa olleen kyselyn. Aivan asennuksen loppuksi on vielä mahdollista määritellä erityinen JSP-sivu raportin katseluun. Tämä siinä tapauksessa jos raporttia ei haluta katsottavaksi JasperServerin oletusnäkyvässä.

Pentahon raportit voidaan julkaista jakelujärjestelmässä suoraan Pentaho Report Designerin kautta. Report Designerissa käyttäjä syöttää jakelujärjestelmän osoitteen ja omat kirjautumistiedot. Kun käyttäjä on saanut kirjaututtua järjestelmään voidaan raportti nimetä, lisätä raportille kuvaus ja valitaan tiedostoformaatti jossa raportti halutaan julkaista. Lopuksi syötetään vielä erillinen julkaisussa tarvittava salasana. Julkaisun salasana on sama kaikilla järjestelmän käyttäjillä. Tämän jälkeen käyttäjän on mahdollista

hypätä suoraan julkaistuun raporttiin. Raportit voidaan asentaa myös jakelujärjestelmässä yksinkertaisesti avaamalla halutut raportit järjestelmään. Asennettavien raporttien tulee olla Pentahon prpt-tiedostoformaattissa.

6.2.3 Raporttien ajaminen

Raportti voidaan ajaa yksinkertaisesti klikkaamalla raportin linkkiä, jolloin sovellus suorittaa tarvittavat kyselyt tietolähteisiin ja raportti näytetään käyttäjälle heti kun tulokset on saatu ja raportti muodostettu. Sekä JasperServerissä että Pentaho BI Serverissä raporttien ajaminen voidaan siirtää suoritettavaksi myös taustalle. JasperServerissä käyttäjä asettaa raportille nimen, tiedostoformaatin ja kansion jossa haluaa lopullisen raportin nähdä. Jos taustalla ajetaan raportti, jonka ajaminen edellyttää käyttäjän syöttämiä parametrejä. Syöttää käyttäjä parametrit samassa yhteydessä kuin määrittää muut yllä mainitut tiedot. Pentaho ei tue raporttien ajamista taustalla, jos ajettavat raportit edellyttävät käyttäjän syöttämiä parametrejä.

6.2.4 Raporttien ajastus

Raporttien ajaminen voidaan myös ajastaa. JasperServer ja Pentaho BI Server tarjoavat kumpikin varsin kattavat mahdollisuudet raporttien ajastamiseen. Raportit voidaan ajastaa ajettavaksi esimerkiksi vain tiettyinä kuukausina, viikonpäivinä tai vain tiettyinä päivinä kuukaudessa. Ajastus voidaan määritellä minuutin tarkkuudella. Kuvio 11 osoittaa kuinka JasperServerillä ajastetaan raportin ajaminen joka maanantaille klo 7.30 ajastuspäivästä alkaen.

Job | Schedule | Parameters | Output

Time Zone

Start Immediately
 On

No Recurrence Simple Recurrence Calendar Recurrence

End date

* Minute(s)
 * Hour(s)

Days Every Day
 Week Days
 Mon Tue Wed Thu Fri Sat Sun

 Month Days

Months All
 Jan Feb Mar Apr May Jun
 Jul Aug Sep Oct Nov Dec

KUVIO 11. Raporttien ajastaminen JasperServerissä.

JasperServerissä sekä taustalla ajettavat, että ajastetut raportit voidaan asettaa lähetettäväksi myös käyttäjän sähköpostiosoitteeseen. Viestiin voidaan aivan normaaliin tapaan lisätä otsikko sekä varsinainen tekstiosuus. Raportti voidaan liittää viestin liitetiedoksi käyttäjän valitsemassa tiedostoformaattissa. JasperServerissä on mahdollista jättää huomioimatta tyhjät raportit. JasperServerissä sähköpostien lähettämisessä tarvittavat tiedot, kuten sähköpostipalvelin, lähettäjän käyttäjätunnus ja salasana sekä viestien välittämisessä käytettävä protokolla ja portti tulee määritellä manuaalisti `js.quartz.properties` -tiedostoon. Pentaho Business Intelligence Serverin vapaasti ladattavissa oleva community editionissa ei ole natiivitukea raporttien lähettämiselle.

6.2.5 Raporttien katselu

Raportteja joille ei ole määritelty erillistä tiedostoformaattia voidaan katsella suoraan jakelujärjestelmien sisällä ilman mitään erillisiä ikkunoita ja ohjelmia.

Mikäli raportissa on useita sivuja ilmestyy raporttinäkymään painikkeet joilla voidaan selata raportin sivuja tai hypätä tietyille sivulle. JasperServerin raporttinäkymästä löytyy kuvakkeet, joista raportti voidaan kääntää eri tiedostoformaateihin ja näin katsoa jollain ulkopuolisella ohjelmalla. JasperServerin tukemat tiedostoformaatit ovat html, pdf, xls, csv, docx, rtf, swf, odt, ods ja xlsx. Pentaho Business Intelligence Serverissä käyttäjä voi valita raportin tiedostoformaatin pudotusvalikosta. Pentahon tukee seuraavia tiedostoformaatteja: html, pdf, xls, csv ja rtf.

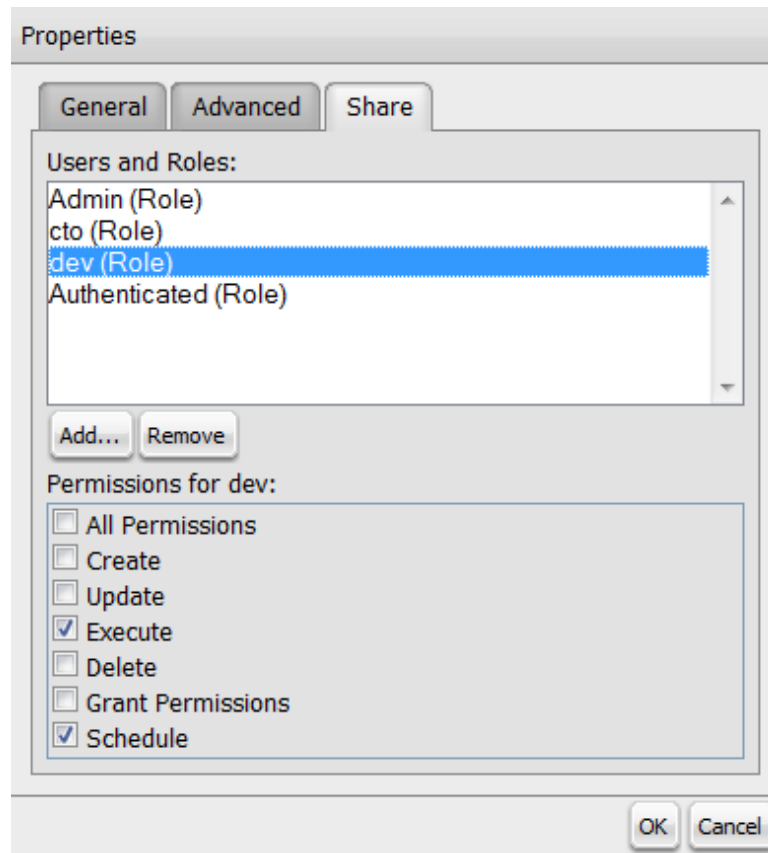
6.2.6 Käyttäjän todentaminen

Sekä JasperServer että Pentaho Business Intelligence Server käyttävät käyttäjien todentamiseen oletuksena järjestelmien omia autentikointimenetelmiä. Kumpikin jakelujärjestelmä hyödyntää kuitenkin Spring Security -ohjelmistokehyksen Acegi Security -lisäosaa, joka puolestaan mahdollistaa muun muassa LDAP, JAAS ja CAS -tekniikoiden käytön käyttäjien tunnistuksessa. LDAPin kautta voidaan hyödyntää muun muassa Windows Active Directory-, Oracle Internet Directory- ja Novell eDirectory - palveluja. Esimerkiksi JasperServerin tiedostoista löytyy kommentoituna valmis pohja LDAP:n käyttöönottoon, joka käyttäjän tulee itse konfiguroida kuntoon ja ottaa käyttöön.

6.2.7 Käyttöoikeudet

Jakelujärjestelmien pääkäyttäjät voivat lisätä järjestelmään rooleja eli käyttäjäryhmiä. Jokaisella järjestelmän käyttäjällä voi olla yksi tai useampi rooli järjestelmässä. Roolien perusteella voidaan jakelujärjestelmän raporteille, raporttihakemistoille tai muille resursseille määrittää käyttöoikeuksia. Käyttöoikeuksia voidaan roolien lisäksi asettaa myös yksittäisille käyttäjille. Käyttöoikeuksilla voidaan esimerkiksi rajata käyttäjältä tai roolilta kokonaan pääsy tiettyyn hakemistoon tai yksittäiseen raporttiin tai myöntää käyttäjälle

vain ajo- ja ajastusoikeus raporttiin (ks. kuvio 12).



KUVIO 12. Käyttöoikeuksien määrittäminen Pentahossa

JasperServerissä hakemistoon annetut oikeudet pätevät oletuksena myös kyseisen hakemiston alihakemistoissa, mutta jokaiselle alihakemistolle voidaan määritellä oikeudet erikseen. JasperServerissä raporttien taustalla ajaminen ja ajastaminen edellyttää että käyttäjällä on kirjoitusoikeus vähintään yhteen hakemistoon, jotta käyttäjän ajamat raportit pystytään luomaan. Ilman kirjoitusoikeutta raporttien taustalla ajaminen tai ajastaminen ei ole mahdollista. Pentaho Business Intelligence Serverissä käyttäjille voidaan antaa erityinen käyttöoikeus, jolla mahdollistetaan raporttien ajastaminen, ilman että käyttäjälle annetaan kirjoitusoikeutta hakemistoon.

6.3 Dokumentaatio

Vapaasti ladattavissa oleville avoimen lähdekoodin sovelluksille on tyypillistä, että sovellusten dokumentointi ja manuaalit ovat maksullisia. Jasperin tuotteista on saatavilla ilman kustannuksia aivan muutama varsin pelkistetty asennus- ja käyttäjäopas. Jaspersoft myy perusteellisempia Ultimate Guide-nimellä kulkevia e-kirjoja erikseen jokaisesta Jasper-tuotteesta. Sovelluskohtaisten oppaiden lisäksi käyttäjän todentamisesta löytyy kokonaan oma e-kirja, jossa käydään yksityiskohtaisesti läpi miten JasperServer konfiguroidaan LDAP- ja CAS-palvelimien kanssa. Oppaat ovat sisällöltään hyvin ajantasalla ja uudet versiot julkaistaan aina suuremman ohjelmistopäivityksen jälkeen. Jaspersoftin virallisten oppaiden lisäksi ainakin Apress ja Packt kustannusyhtiöt ovat julkaisseet muutaman Jasper-sovelluksia käsittelevän kirjan. Kirjojen lisäksi JasperForgen sivulla on nähtävissä aivan muutamia tutorial-artikkeleita ja muita videoita sovellusten perusominaisuuksien käytöstä. Artikkelit ja videot eivät kuitenkaan mene syvälle sovelluksiin ja pelkästään niiden perusteella suurempien muutosten ja konfigurointien tekeminen sovelluksiin ei onnistu.

Pentahon sivuilta löytyy sekä Report Designeria että Business Intelligence Serveriä käsittelevät wikit. Wikit eivät kuitenkaan ole lainkaan ajantasalla vaan osa wiki-sivuista käsittelee vielä Pentahon sovellusten 1.x tai 2.x -versioita ja monet wiki-sivuista ovat siitä huolimatta keskeneräisiä. Pentahon wiki on täysin Pentahon käyttäjien ylläpitämä. Ainoastaan Pentahon kaupallisen Enterprise Editionin käyttäjät saavat käsiinsä ammattimaisesti kirjoitetun ja ylläpidetyn dokumentaation. Tämä Pentahon virallinen dokumentaatio kattaa vasta sovellusten 2.0 -versiot. Kustannusyhtiö Packt on julkaissut syksyllä 2009 yhden Pentaho-raportointia käsittelevän kirjan, jonka lisäksi Wileylta on tullut yksi koko Pentaho-tuoteperhettä koskeva kirja.

7 JOHTOPÄÄTÖKSET JA POHDINTA

Sekä Jasperin että Pentahon avoimen lähdekoodin raportointivälineet olivat ominaisuuksiltaan erittäin monipuolisia kokonaisuuksia. Kummastakin tutkitusta raportointisovelluksesta löytyi tuet useille kymmenille eri tietolähteille ja tietojen käsitteleminen mahdollistettiin suorien SQL-kyselyiden lisäksi myös proseduureilla. JasperReportsia hyödyntävä iReport suoriutui kaikista perusraportoinnille asetetuista vaatimuksista erinomaisesti. Pentaho Report Designerin ainoaksi merkittävästi puutteelliseksi ominaisuudeksi osoittautui ristiintaulukointi. Pentaho Report Designerin ristiintaulukointi toimii tällä hetkellä hyvin alkeellisella tasolla ja jopa sen peruskäyttäminen on vaivalloista ja hidasta.

Tutkituissa jakelujärjestelmissä käyttäjien hallintaan on monipuoliset mahdollisuudet, kummankin salliessa käyttäjien todentamisen ja käyttöoikeuksien jakamisen yleisimpien käytössä olevien hakemistopalvelujen kanssa. JasperServer ja Pentaho Business Intelligence Server mahdollistavat raporttien ajastamisen, mutta vain ensiksi mainitusta löytyy sisäänrakennettu tuki raporttien lähettämiseen. Raporttien lähettäminen voi olla mahdollista toteuttaa myös Pentahossa, mutta tämä vaatii järjestelmän toimittajalta hyvin huomattavan määrän lisätyötä. Pentahon ratkaisuihin tehtävien lisäyksien ja muutosten toteuttamista vaikeuttaa merkittävästi virallisen dokumentoinnin puute. Kaupalliseen versioon on tarjolla dokumentaatiota, mutta sen sisällöstä ja ajantasaisuudesta ei ole selvyyttä.

Jasperin ja Pentahon raportointivälineet ovat yllämainittuja puutteita lukuunottamatta erittäin samankaltaisia ominaisuuksiensa puolesta. Sen sijaan raportointivälineiden yleisessä toteutuksessa on eroavaisuuksia. Vaikka

tutkimuksella ei selvitetty työvälineiden käytettävyyttä niin sovelluksia havainnoidessa huomasi, että Jasperin sovellukset ovat paremmin viimeisteltynä ja selkeämpinä käyttää. iReportissa monet yksinkertaisistakin perustoiminnoista on hyvin havainnollistettuja kuvin tai kuvakkein, eikä käyttäjän tarvitse ensimmäiselläkään käyttökerralla arvailla tai testaila kuinka eri toimintoja käytetään. Pentaho Report Designerissa ei ole lainkaan ohjattuja toimintoja ja ominaisuuksien havainnollistaminen on lähes olematonta. Pentaho Report Designerin käyttöliittymässä on esimerkiksi sellainen erikoisuus, että raporttiin lisättyjä raporttielementtejä ei voida siirtää raportin eri osioiden välillä helposti.

Jasperin raportointivälineet vastaavat tutkittujen ominaisuuksien, teknologisten ratkaisujen ja yleisen toteutuksen puolesta erinomaisesti kaupallisia tuotteita. Pentahon raportointivälineet ovat esitetyistä puutteista huolimatta hyvin laadukkaita avoimen lähdekoodin sovelluksia. Ominaisuuksien osalta sekä Jasperin että Pentahon tuotteilla pystytään korvaamaan kaupalliset tuotteet vähintään perinteisimmässä business intelligence -ratkaisussa.

Opinnäytetyön tavoitteena oli selvittää kuinka hyvin avoimen lähdekoodin business intelligence -raportointivälineet vastaavat ominaisuuksiltaan ja teknologisilta ratkaisuiltaan kaupallisia business intelligence tuotteita. Työ onnistui odotusteni mukaisesti ja sain selvitettyä tutkittavina olleista raportointivälineistä kaikki toimeksiantajan niille asettamat vaatimukset. Raportointivälineistä tuotiin esiin lisäksi niiden merkittävimpiä puutteita ja välineitä verrattiin keskenään.

Työn lopputuloksena saatiin yksityiskohtainen raportti avoimen lähdekoodin business intelligence -raportointivälineiden ominaisuuksista. Koska työssä on tutkittu juuri toimeksiantajan erikseen asettamia vaatimuksia, vastaa työ

ensisijaisesti toimeksiantajan tarpeeseen, mutta myös muut business intelligence -ratkaisujen parissa toimivat yritykset voivat hyötyä tutkimuksen tuloksista. Raportin avulla työn toimeksiantaja saa hyvän käsityksen laadukkaimmista avoimen lähdekoodin raportointivälineistä ja pystyy kehittämään omaa liiketoimintaansa avoimen lähdekoodin business intelligence -tuotteilla.

Tutkimusprosessiin ei kuulunut raportointivälineiden toiminta ja toimivuus todellisessa käyttöympäristössä, oikeilla tiedoilla ja käyttäjillä. Tämän tutkimuksen suorana jatkotutkimuksena voisikin olla tässä tutkimuksessa tutkittujen työvälineiden soveltaminen todellisuudessa. Jatkotutkimuksena voitaisiin tehdä esimerkiksi kehitystyö, jossa yritykselle tai organisaatiolle toteutetaan business intelligence -ratkaisu avoimen lähdekoodin tuotteilla. Toinen mahdollisuus voisi olla suorittaa kyselytutkimus yritykselle, joka on ottanut käyttöönsä avoimen lähdekoodin ratkaisun tai jopa siirtynyt kaupallisesta ratkaisusta avoimeen lähdekoodiin. Kumpikin jatkotutkimus olisi aiheena erittäin haastava, mutta oppimisprosessina erittäin antoisa ja osaamista kehittävä, aivan kuten tämäkin tutkimus.

LÄHTEET

About Pentaho. 2009. Pentaho. Viitattu 29.4.2010.

<http://www.pentaho.com/about/>.

Abukari, K. & Jog, V. 2003. Business intelligence in action. CMA Management, March 2003, 15 - 18.

Azvine, B., Cui, Z. & Nauck, D. D. 2005. Towards real-time business intelligence. BT Technology Journal, 23, 3.

Bitterer, A. 2008. Who's Who in Open-Source Business Intelligence. Gartner.

Tutkimusraportti, julk. 16.4.2008. Viitattu 1.7.2009.

http://www.stratebi.es/todobi/may08/whos_who_in_opensource_busin_156326.pdf.

Bouman, R. & van Dongen, J. 2009. Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL. Indianapolis: Wiley Publishing, Inc.

Business Intelligence. 2009. Lingaro. Viitattu 2.11.2009.

<http://www.lingaro.com/lingaro/competencies/business-intelligence/>.

Gantz, J. & Reinsel, D. 2009. As the Economy Contracts, the Digital Universe Expands. IDC Multimedia White Paper, julk. toukokuu 2009. Viitattu 1.7.2009.

http://idcdocserv.com/EMC_MMWP_Digital_Universe.

Grimes, S. 2008. BI at 50 Turns Back to the Future. Intelligent Enterprise.

Julkaistu 2.11.2008. Viitattu 2.9.2009.

<http://www.intelligententerprise.com/showArticle.jhtml;jsessionid=4YFOIQYBF0MENQE1GHOSKHWATMY32JVN?articleID=211900005>.

GNU Affero General Public Licence v3. 2009. Open Source Initiative. Viitattu 27.11.2009. <http://opensource.org/licenses/agpl-v3.html>

Hirsjärvi, S., Remes, P. & Sajavaara, P. 2009. Tutki ja kirjoita. 15. p. Helsinki: Tammi.

Howson, C. 2008. Successful Business Intelligence: Secrets to Making BI a Killer App. New York: McGraw-Hill.

Hovi, A., Hervonen, H. & Koistinen, H. 2009. Tietovarastot ja Business Intelligence. Jyväskylä: Docendo.

Ingo, H. 2005. Avoin elämä. <http://www.avoinelama.fi/>.

JasperForge. 2009. Jaspersoft Corporation. Viitattu 27.11.2009. <http://jasperforge.org/>

Järvinen, P. & Järvinen, A. 2004. Tutkimustyön metodeista. Tampere: Opinpajan kirja.

Laurent, A. M. S. 2004. Understanding Open Source and Free Software Licensing. California: O'Reilly.

Moody, G. 2001. Kapinakoodi: Linus Torvalds ja vapaan ohjelmoinnin vallankumous. Helsinki: Tammi.

Moore, S. 2009. Business Intelligence Ranked Top Technology Priority by CIOs for Fourth Year in a Row. Gartner. Lehdistötiedote yhtiön verkkosivuilla, julk. 18.2.2009. Viitattu 1.7.2009. <http://www.gartner.com/it/page.jsp?id=888412>.

OLAP Council. 1997. OLAP Council White Paper. Viitattu 12.10.2009. <http://www.olapcouncil.org/research/whtpaply.htm>.

Open Source Initiative. Viitattu 18.8.2009. <http://www.opensource.org/>.

Peck, G. 2008. Crystal Reports 2008: The Complete Reference. New York: McGraw-Hill.

Pentaho. 2009. Viitattu 29.4.2010. <http://www.pentaho.com/>.

Pentaho - Business Intelligence. 2010. SourceForge. Viitattu 29.4.2010. <http://sourceforge.net/projects/pentaho/>

Pettey, C. & Stevens, H. 2009. Gartner Says Worldwide Business Intelligence, Analytics and Performance Management Grew 22 Percent in 2008. Gartner. Lehdistöiedote yhtiön verkkosivulla, julk. 12.6.2009. Viitattu 2.9.2009. <http://www.gartner.com/it/page.jsp?id=1017812>.

Power, D. J. 2007. A Brief History of Decision Support Systems. <http://dssresources.com/history/dsshhistory.html>.

SourceForge. 2010. Viitattu 29.4.2010. <http://sourceforge.net/softwaremap/>

The Free Software Definition. 2009. Free Software Foundation. Päivitetty 21.8.2009. Viitattu 21.8.2009. <http://www.fsf.org/licensing/essays/free-sw.html>.

The Open Source Definition. 2009. Open Source Initiative. Viitattu 18.8.2009. <http://opensource.org/docs/osd>.

Thierauf, R. J. 2001. Effective Business Intelligence Systems. Westport: Quorum Books.

Thomas Jr., J. H. 2001. Business Intelligence - Why? eAI Journal, july 2001. 47-49.