



Laboratorion raportointitarpeiden kartoitus ja toteutuksen suunnittelu

Siru Hämäläinen

2021 Laurea





Laurea-ammattikorkeakoulu

Laboratorion raportointitarpeiden kartoitus ja toteutuksen suunnittelu

Siru Hämäläinen
Tietojenkäsittely
Opinnäytetyö
Syyskuu, 2021

Siru Hämäläinen

Laboratorion raportointitarpeiden kartoitus ja toteutuksen suunnittelu

Vuosi

2021

Sivumäärä

36

Opinnäytetyön tavoitteena oli selvittää minkälaisia raportteja Vita Laboratoriot Oy:n työntekijät tarvitsevat työssään. Toimeksiantona oli selvittää, miten raportteihin saatava data saadaan käytössä olevasta Molis-laboratoriojärjestelmästä ja minkälaisia vaatimuksia järjestelmät sekä raportointitarpeet asettavat tulevalle tietovarastolle. Raporttien jako- ja visualisointiohjelmistoksi oli yrityksessä valittu Microsoftin kehittämä Power BI.

Tässä opinnäytetyössä selvitettiin teemahaastatteluilla minkälaisia raportteja Vita laboratoriot Oy:n markkinointiosaston, tuotannon sekä johdon henkilökunta tarvitsee työnsä tueksi. Haastatteluiden tulosten perusteella muodostettiin kahdeksan erilaista raporttimallia Power BI ohjelmistolla. Haastatteluissa selvisi, että henkilökunta on hyvin tietoinen raportoinnin hyödyistä ja haluaisi hyödyntää kerättyä dataa paremmin työssään. Esiin nousi myös tyytymättömyys tämänhetkiseen raportointijärjestelmään ja raportteja toivottiin helpommin ja nopeammin saataville. Raportteja tarvitaan vastatuista tutkimusmääristä jaoteltuina tutkimusryhmittäin, analysoitavina ja toimiala- sekä yrityskohtaisesti. Lisäksi henkilökunta tarvitsee raportteja sisään kirjattujen näytteiden määrästä, tutkimusten vastausviiveistä sekä keskimääräisistä numeerisista vastauksista.

Kehittämisen ja jatkotoimenpiteenä laboratoriojärjestelmään tulisi määritellä erilaiset vakio- vastaukset näytepoikkeamille, jotta niitä voidaan raportoida luotettavasti. Lisäksi laskutusta voidaan kehittää hyödyntämällä tulevan tietovaraston toiminnallisuuksia ja vähentää manuaalista työtä. Toimeksiantajayrityksen käytössä olevien muiden laboratoriojärjestelmien liittäminen tietovarastoon on suuri projekti ja sitä tulisi punnita sekä taloudellisista, että ajankäytöllisistä näkökulmista.

Siru Hämäläinen

Mapping of Laboratory Reporting Needs And Planning the Implementation

Year 2021 Pages 36

The objective of this Bachelor's thesis was to investigate what kind of reports the Vita Laboratoriot Ltd. personnel needs to support their work. The purpose was to find out how to get the needed data from Molis laboratory system and what kind of requirements are set to the data warehouse by the laboratory systems. The software used to visualize and share the reports was Power BI.

In this study, the needs of the marketing department, production and management were mapped out through semi-structured interviews. Based on the results eight report models were created with Power BI. According to the results the personnel is aware of the ways reports and knowledge-based decisions can be utilized in their work, and they are discontented with the current state of reporting. The reports need to be available more easily. The reports are needed on answered laboratory tests broken down by test groups, industry, analyser and companies. In addition, reports are needed on tubes registered to the laboratory system, test result delay and average test results.

As a development proposal it would be necessary to define standard results to the laboratory system to tests with sample error. After it is done, it is possible to report the errors reliably. In addition, invoicing can be developed to more automated through the features offered by the data warehouse. It is possible to connect all the laboratory systems used in Vita Laboratoriot Ltd. into the same data warehouse, but this issue should be considered first from economical and time-driven perspectives.

Keywords: Power BI, business intelligence, ETL-process, data warehouse, knowledge-based management

Sisälllys

1	Johdanto.....	7
2	Työn lähtökohdat.....	7
2.1	Kehittämistavoitteet.....	8
2.2	Aihealueen rajaus	8
2.3	Keskeiset käsitteet.....	9
3	Tietovarastointi ja BI	10
3.1	Tiedon laatu ja sen ongelmat	10
3.2	ETL-prosessi.....	11
3.3	Tietovarastointi terveydenhuoltoalalla	12
3.4	Tiedolla johtaminen	13
3.5	Datan visualisointi Power BI:llä	13
4	Kehittämismenetelmät	15
4.1	Tutkimuksen luotettavuuden arviointi	15
4.2	Eettisyys.....	15
4.3	Teemahaastattelu.....	16
4.4	Haastattelujen analyysi	17
5	Case: BI raportointi Vita Laboratoriot Oy:ssä	17
5.1	Henkilökunnan haastattelut	18
5.2	Raportoinnin lähtötilanne.....	22
5.3	Tiedonhaku Moliksesta	22
5.4	Datan laadun muokkaus ja ETL-prosessi	24
5.5	Tietovarastoratkaisu.....	26
5.6	Power BI raportit	27
6	Yhteenveto	31
7	Jatkokehittämissuositukset	32
	Lähteet.....	33
	Kuviot	36
	Taulukot	36

1 Johdanto

Tämän opinnäytetyön toimeksiantona on tutkia ja selvittää minkälaisia raportteja ja tilastoja Vita Laboratoriot Oy:n henkilökunta tarvitsee työnsä tueksi ja tehdäksään tietoon perustuvia päätöksiä. Tavoitteena on myös selvittää, miten raportit voidaan muodostaa ja minkälainen tietovarastoratkaisu sopii yrityksen tarpeisiin parhaiten. Opinnäytetyön toimeksiantaja on Vita Laboratoriot Oy.

Tiedolla johtaminen ja tiedon hyödyntäminen on nykYTEKNIKAN avulla yhä helpompaa. Kilpailu on monella alalla kovaa ja tietoon perustuvat päätökset voivat olla ratkaisevassa osassa yrityksen pärjäämisessä markkinoilla. Usein tietoteknisiin ratkaisuihin ei haluta panostaa rahallisesti ja aiempaan vanhaan raportointitapaan on jo totuttu. Rahallinen panostus kuitenkin maksaa itsensä pian takaisin sekä ajansäästönä, että parempien tietoon perustuvien päätösten muodossa. (Miller, Dolan, Crowson & Stout 2002, 74.)

Vita Laboratoriot Oy on vuonna 1994 perustettu suomalainen yritys, joka tuottaa pääasiassa keskuslaboratoriopalveluita muille yrityksille. Asiakkaina on yksityisiä yrityksiä, lääkäriasemia sekä julkisen puolen toimijoita kuten sairaanhoitopiirejä. Vuonna 2012 Vita Laboratoriot laajensi toimintaansa ympäristölaboratorioksi ja tarjoaa nykyään myös sisäilma, vesi-, elintarvike-, hygieniä ja asbestitutkimuksia. (Hurme, 2021.) Vita Laboratoriot Oy:n liikevaihto oli vuonna 2020 19,9 miljoonaa euroa ja se työllisti 110 henkilöä. Liikevaihto nousi edellisvuoteen verrattuna 206,1 prosenttia. (Vita Laboratoriot 2020.)

Vita Laboratoriot Oy:n toiminta on kasvanut voimakkaasti viimeisen vuoden aikana ja tarve tehokkaammalle myynti- ja tutkimusmäärien raportoinnille on kasvanut laajenemisen myötä. Entinen tapa raportoida ja käsitellä myynti- ja tutkimusdataa ei ole enää tarkoituksen mukainen ja siihen toivotaan tehokkaampaa toimintatapaa. Opinnäytetyö on osa yrityksen projektia, jonka tarkoituksena on saada kaikkien käytössä olevien laboratoriojärjestelmien myynti-, tutkimus- ja hintatilastot samaan ohjelmistoon. Raporttien muodostamiseen ja visualisointiin on valittu ohjelmistoksi Microsoftin Power BI. (Hurme 2021.)

2 Työn lähtökohdat

Vita Laboratorioilla on käytössään useita erilaisia laboratoriotietojärjestelmiä. Kliinisellä puolella on käytössä Molis (Modular Open Laboratory Information System) ja ympäristöpuolen laboratoriossa käytetään BroadSight Lims -ohjelmaa. Lisäksi patologian osasto käyttää QPati -ohjelmistoa ja yrityksellä on myös verkkokauppa kuluttajille. Laboratoriojärjestelmillä ei ole

tällä hetkellä yhtenäistä raportointikanavaa, ja tilastojen kokonaiskuvan hahmottaminen on vaikeaa. (Hurme 2021.)

Opinnäytetyö on osa yrityksen projektia, jonka tavoitteena on ottaa käyttöön uusi yhtenäinen tutkimustilastojen tietovarastointi- ja raportointijärjestelmä. Uutta järjestelmää tulevat käyttämään Vita Laboratoriot Oy:n markkinointiosasto, johdon henkilöstö sekä tuotannosta vastaavat asiantuntijat. Lisäksi joitain raportteja halutaan jakaa koko henkilöstölle. (Hurme 2021.)

Uuden raportointijärjestelmän hyödyt näkyvät koko yrityksessä. Ajansäästö näkyy myös IT-osastolla, koska datahakuja ei tarvitse tehdä erillisinä hakuina tietokannasta. IT -osaston henkilökunta on tehnyt datahakuja muun henkilöstön pyynnöstä muutamia kertoja kuukaudessa, ja raporttien käyttäjät ovat itse muokanneet haetun datan haluamaansa muotoon. Projektin valmistuttua, datan hakuun ja muokkaukseen käytettävä aika vähenee ja raportteja pystytään hyödyntämään huomattavasti tehokkaammin ja useammin. (Päivinen 2021.)

2.1 Kehittämistavoitteet

Tämän opinnäytetyön tavoitteena on selvittää, minkälaisia raportteja yrityksen henkilökunta tarvitsee työssään, ja miten ne voidaan toteuttaa valitulla Power BI -ohjelmistolla. Yrityksen henkilökunnan tarpeet raportoinnille selvitetään haastatteluilla ja lisäksi selvitetään mitä tietoja laboratoriojärjestelmästä haetaan ja miten niitä täytyy muokata, jotta tarvittavat raportit pystytään muodostamaan. Lisäksi kartoitetaan minkälainen tietovarastoratkaisu olisi yrityksen tarpeisiin sopiva. Tietovarastoon pitäisi pystyä tulevaisuudessa yhdistämään kaikki yrityksen käytössä olevat laboratoriojärjestelmät. (Päivinen 2021.)

Tutkimuskysymykset:

- Mitä tarpeita johdolla, markkinointiosastolla ja tuotannolla on raporttien suhteen?
- Mitä tietoja Moliksesta haetaan ja miten niitä muokataan tarvittavien raporttien muodostamiseen?
- Mitä vaatimuksia tietovarastolla on yrityksen tarpeisiin nähden?

2.2 Aihealueen rajaus

Työssä selvitetään Molis -laboratoriojärjestelmän tietokantarakennetta ja sieltä tarvittava data sekä kartoitetaan olemassa olevan datan laatua. Lisäksi selvitetään henkilökunnan tarpeet ja vaatimukset raporteille. Muiden käytössä olevien laboratoriojärjestelmien kohdalla selvitetään tekninen yhteensopivuus, mutta ei oteta kantaa siihen mitä tietoja järjestelmistä kannattaa hakea ja miten se toteutetaan. Opinnäytetyössä ei oteta kantaa

tietovarastoratkaisun kaupallisiin vaihtoehtoihin, eikä kartoiteta muita raportointityökaluja Power BI ohjelmiston lisäksi.

Opinnäytetyön teoria osuudessa käsitellään tietovarastoinnin ETL-prosessia sekä datan laadun muokkausta. Käytetyt laadullisen tutkimuksen menetelmät esitellään tarkemmin vain käytetyt menetelmät.

2.3 Keskeiset käsitteet

Business Intelligence tarkoittaa liiketoiminnan hallintaa ja suunnittelua erilaisten sovellusten ja teknologioiden avulla (Törmänen 1999, 201). Sille ei ole vakiintunutta suomenkielistä vastinetta, mutta esimerkiksi termi liiketoimintatiedonhallinta kuvaa käsitettä. Sovellusten avulla kerätyn informaation avulla henkilöstö ja johto pystyvät tekemään informaatioon perustuvia päätöksiä. (Hovi, Hervonen & Koistinen 2009, 74.)

ETL-prosessi tarkoittaa prosessia, jossa data haetaan (engl. extract) eri tietolähteistä, se muokataan yhtenäiseksi (engl. transform) ja lopuksi data ladataan tietovarastoon (engl. load). (Silvers 2011, luku 13.)

Metadata on tietoon linkitettyä tietoa eli esimerkiksi tietoa siitä, milloin ja mistä data on ladattu ja miten sitä on muokattu (Kimball, Ross, Becker, Mundy, & Thornthwaite 2008, luku 10).

Molis eli Modular Open Laboratory Information System, on Vita Laboratoriossa käytettävä kliinisen puolen laboratoriojärjestelmä. Ohjelmasta on sähköinen yhteys välityspalvelimien kautta asiakasyritysten laboratoriojärjestelmiin. Ohjelmaa pystyy muokkaamaan yrityksen tarpeisiin ohjelmassa käytettävällä MPL-ohjelmointikielellä. (CGM 2021.)

Power BI on Microsoft Corporationin kehittämä ohjelmisto, joka on tarkoitettu datan visualisointiin ja käsittelyyn. Sen avulla voidaan tehdä interaktiivisia raportteja, jotka ovat helposti käyttäjän muokattavissa. (Microsoft 2021a.)

Relaatiotietokanta on yleensä SQL-pohjainen datanvarastointijärjestelmä, jossa tiedot on tallennettu erilaisiin tauluihin. Taulujen välille on rakennettu erilaisia riippuvuussuhteita ja linkkejä. Tauluissa olevia tietoja kutsutaan attribuuteiksi. (Hovi, Ylinen & Koistinen 2001, 57).

Tietovarasto on prosessi, jolla kerätään ja hallinnoidaan eri lähteistä kerättyä tietoa ja jalostetaan, yhdistellään ja muokataan tieto paremmin hyödynnettävään muotoon. Tietovarasto on usein BI-järjestelmän ydin. Tietovarastoon ladataan dataa useasta eri lähteestä, esimerkiksi relaatiotietokannoista. Yhdistämällä kaikki tiedot yhteen paikkaan, organisaatio voi analysoida paremmin saatavilla olevaa tietoa ja etsiä datasta erilaisia malleja ja suuntauksia. (Hovi ym. 2009, 14-16.)

Työalueen (engl. staging area) tehtävä on avustaa ETL-prosessia ja sen sisällä datan rakenne on sama kuin tietolähteessä. Data ladataan operatiivisista järjestelmistä ensimmäiseksi työalueelle. (Hovi ym. 2009, 190.)

3 Tietovarastointi ja BI

Tietovarastoprojektin tavoitteiden tulisi aina olla lähtöisin liiketoiminnan tavoitteista ja tarpeista (Hovi ym. 2001, 19). Kaikkiin tiedonsaantiongelmiin tietovarastointi ei ole oikea ratkaisu, vaan niitä täytyy lähteä ratkaisemaan organisaatiomuutosten tai tehtävien organisoinnin kautta. Tietovarastoinnin suunnitteluun kannattaa varata hyvin aikaa, jotta siitä saatava hyöty vastaa asetettuja tavoitteita. Lisäksi projektilla tulee olla vahva johdon tuki. Tietovarastoinnin ja BI raportoinnin perusrakenne on kuvattu kuviossa 1. Tieto haetaan eri tietolähteistä. Tiedon muokkauksen jälkeen tieto ladataan tietovarastoon, josta se siirtyy datan visualisointi ohjelmaan (Power BI). Loppukäyttäjät hyödyntävät visualisoituja raportteja työssään. (Törmänen 2017,10.)



Kuvio 1: Tietovarastointi ja BI raportointiprosessi

3.1 Tiedon laatu ja sen ongelmat

Tietovarastoon haettavien tietojen laatuun ja ylläpitoon kannattaa panostaa projektin alusta lähtien. Jos tietovarastosta saatavan tiedon laatu on huono, sen hyödynnettävyys ja käyttöaste tulevat olemaan matalat. (Hovi ym. 2001, 34; Törmänen 2017, 32.) Tiedon laadulle tulisi asettaa yrityksen sisällä tavoitteet ja tehdä laatustrategia. Täydellisesti oikeat tiedot eivät välttämättä ole saavutettavissa vaan erityyppisille tiedoille voidaan sallia virheitä tiettyjen raja-arvojen sisällä. (Hovi ym. 2009, 69.) Tietovaraston ja raporttien suunnittelussa on tärkeää kartoittaa käyttäjien tarpeet raporttien ja haettavan tiedon suhteen. Käyttäjien tarpeet tulee huomioida myös valittavien raportti- ja hakutyökalujen valinnassa. (Törmänen 2017, 32–33.)

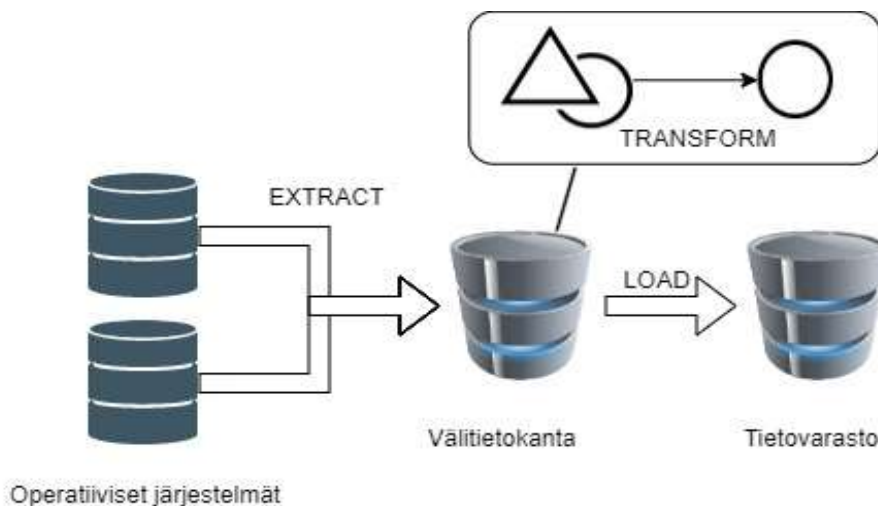
Kerättävän datan ongelmia voivat olla tietotyyppien erilaisuus, taulujen sisäisten viittausten puuttuminen, perusavaimet ovat erilaiset eri järjestelmissä tai tietokenttiin operatiivisessa järjestelmässä syötetty vääränlainen informaatio (Hovi ym. 2001, 37). Operatiiviseen järjestelmään väärin syötetty data tulisi korjata ja sen jälkeen ladata uudelleen tietovarastoon. Datat laadun korjaamiseen tulisikin varata runsaasti aikaa. (Hovi ym. 2001, 38.)

3.2 ETL-prosessi

Laadukas ETL-prosessi vaikuttaa merkittävästi tiedon laatuun ja siten koko projektin onnistumiseen. (Souibgui, Atigui, Zammali, Cherfi & Yahia 2019, 677.) Prosessissa määritellään ohjelmallisesti hyväksyttävät ja hylättävät data-arvot. Tätä varten on tiedettävä, mitä hyväksyttävät tiedot ovat, ja ne on voitava määritellä ohjelmallisesti. Yleensä tietojen laadun parantaminen on iteratiivinen toimenpide ja siihen tulee varata riittävästi aikaa. Aktiivinen ja toimiva tiedon laadun varmistava ohjelma parantaa tietovarastoprojektin onnistumista merkittävästi. (Silvers 2011, luku 1.) ETL-prosessin ja tiedon laadun parantamiseen on olemassa paljon erilaisia työkaluja ja ohjelmistoja. Niiden kyky muokata dataa on kuitenkin rajallinen. Ohjelmistojen etuna on esimerkiksi laadukkaan metadatan kerääminen ja sen parempi hyödyntäminen. (Khan, Ehsan, Mirza & Sarwan 2012, 681.)

Kuviossa 2 on kuvattu ETL-prosessin eteneminen. Extract -vaiheessa tieto poimitaan erityyppisistä järjestelmistä. Useimmiten järjestelmät ovat operatiivisia järjestelmiä. Tieto siirretään välitietokantaan eli niin kutsutulle työalueelle (engl. staging area). Tiedon siirrossa tulee huomioida se, että lataus hidastaa operatiivisen järjestelmän toimintaa ja se tulisikin toteuttaa siten että operatiivista järjestelmää rasitetaan mahdollisimman vähän. (Hovi ym. 2001, 78.)

Tiedon muuntovaiheessa (engl. transform) turhaa tietoa poistetaan, dataa yhdistellään, siivotaan ja tiedoissa olevia virheitä korjataan. Tämä vaihe on yksi kriittisimmistä vaiheista tietovarastointiprojektin onnistumisen kannalta. (Kimball ym. 2008, luku 10.) Tärkeitä tarkistuksia muuntovaiheessa ovat profilointi, yhdistäminen ja monitorointi. Profiloinnilla tarkistetaan tiedon laatua esimerkiksi tarkistamalla, kuinka montaa erilaista sukupuolikoodia datasta löytyy. Yhdistämisellä tarkoitetaan samaa tarkoittavien koodien etsimistä, vertailua sekä integrointia. Monitoroinnilla tarkoitetaan datan laadun seuraamista koko prosessin ajan. (Hovi ym. 2001, 78.)



Kuvio 2: ETL-prosessin eteneminen

Load - vaiheessa muokattu ja siivottu data ladataan varsinaiseen tietovarastoon. Lataus voidaan tehdä käyttäen latausohjelmia tai ETL-työvälineitä. Useimmiten tiedot lisätään vanhojen tietojen perään ja lataus toteutetaan öisin, jolloin se häiritsee mahdollisimman vähän lähdejärjestelmien ja yrityksen toimintaa. Ajojen seurantaan käytetään erilaisia logeja, joista nähdään mm. siirron kesto ja siirretyn datan määrä. (Hovi ym. 2009, 58.)

3.3 Tietovarastointi terveydenhuoltoalalla

Laboratoriotietovaraston kehittäminen mahdollistaa helpon pääsyn tarkkoihin ja asiaankuuluviin lääketieteellisiin- sekä yritystietoihin, joiden avulla asiantuntijat saavat enemmän tietoa päätöksien tueksi paitsi potilaiden diagnosoinnissa ja hoidossa, myös päivittäisissä toiminnoissa ja strategisessa suunnittelussa. Tietovaraston tavoitteena on tuoda organisaation tiedot helposti saataville ja muokata tiedoista johdonmukaiset, jotta tiedon perusteella voidaan tehdä päätöksiä. Tietovaraston pitää olla mukautuva, joustava ja turvallinen tietolähde. (Miller ym. 2002, 74.)

Terveydenhuoltoalalla tietolähteet ovat usein erilaisia kuin muilla aloilla ja se asettaa tiettyjä vaatimuksia tietovarastolle. Tietolähteet vaihtelevat laboratoriovastauksista, patologian laitteisiin sekä vieritestilaitteista kansallisiin rekistereihin. Tiedon moninaisuuden takia, sen hyödyntäminen on usein vaikeaa. Muilla aloilla käytetyt tietovarastoratkaisut eivät ole aina sovellettavissa terveydenhuoltoalalle. Tiedon tarve on sekä hallinnollista että kliinistä. Terveydenhuoltoalalla käsitellään potilaiden arkaluonteista tietoa, joten tietosuojaan on kiinnitettävä erityistä huomiota. (George, Kumar & Kumar 2015.)

Suunnitteluvaiheessa tulee kiinnittää huomiota tietolähteiden valintaan. Kaikkea tietoa kaikista lähteistä ei tarvita eikä pystytä ottamaan mukaan. Usein terveydenhuollon

tietovarastoissa käytetään tähti-, sekä lumihitalemallia. Terveystietojen datan kaksi näkökulmaa (hallinnollinen ja kliininen) määräävät suunniteltavan arkkitehtuurin rakenteen. Nämä kaksi puolta eivät sulje toisiaan pois vaan kulkevat rinnakkain koko projektin läpi. Työalueen (staging area) tietoturvaan ja siellä käsiteltävään dataan tulee kiinnittää erityistä huomiota. (George ym. 2015.)

3.4 Tiedolla johtaminen

Johtamisen peruserä on, että prosessin mittaus johtaa hallintaan, joka puolestaan johtaa kehittämiseen. Tämän periaatteen noudattaminen vaatii sitä, että tarkat ja olennaiset tiedot ovat oikea-aikaisesti päätäjien saatavilla. (Miller ym. 2002, 74.) Tietovaraston sisältämät tiedot ovat vain numeroarvoja, ellei niitä analysoida ja esitetä ymmärrettävästi käyttäjille. Termi tiedonlouhinta perustuu oletukseen, että valtava määrä dataa kerätään ja sen suhteita vertaillaan ja mallinnetaan suhteessa toisiinsa, aiempaan tietoon tai muusta lähteestä kerättyyn dataan. Tehtyjä mallinnuksia voidaan käyttää ennusteiden tekemiseen. (Miller ym. 2002, 75.)

Yritys tarvitsee tietoa ymmärtääkseen toimintaansa, asiakkaitaan, kilpailijoitaan, yhteistyökumppaneitaan sekä työntekijöitään ja heidän toimintaansa. Tietoon perustuva johtaminen on nykyaikaista ja tiedon avulla tehdyt päätökset ovat perusteltuja ja johdonmukaisia. (Sherman 2014, luku 1.)

Yritykselle tärkeimmästä tiedosta käytetään termiä master data. Se on yritykselle kriittisen tärkeää tietoa, jota koko organisaatio käyttää. Väre (2019, Luku 1) jakaa master datan kolmeen eri tasoon. Enterprise master data on koko organisaation hyödyntämää dataa. Toiminto- tai tuotealuekohtainen master data on tietyn toiminnan tai osa-alueen käyttämää dataa. Järjestelmä- tai prosessikohtainen data eli se data, jota järjestelmä vaatii käyttämään, on oleellista erityisesti tietylle prosessille tai järjestelmän osalle.

Yrityksen toiminnasta saadun hyvin strukturoidun ja raportoidun tiedon perusteella yrityksen johto voi tehdä perusteltuja päätöksiä liittyen esimerkiksi investointeihin sekä toimintavaihtoehtoihin. Laadukkaan ja helposti saatavilla olevan tiedon perusteella pystytään myös arvioimaan esimerkiksi kilpailijoiden tekemien toimenpiteiden vaikutuksia oman yrityksen liiketoimintaan sekä ennustaa ympäristössä tapahtuvien muutoksen vaikutuksia yrityksen toimintaan. (Törmänen 1999, 36.)

3.5 Datan visualisointi Power BI:llä

Microsoftin kehittämän Power BI ohjelmiston avulla on mahdollista luoda interaktiivisia raportteja ja datan visualisointeja. Ohjelmistoon kuuluu työpöytäsovellus Power BI Desktop, jolla voidaan muodostaa yhteys erilaisiin datalähteisiin. Sillä myös muokataan dataa sekä

muodostetaan raportteja sekä erilaisia datan visualisointeja. Muita ohjelmiston osia ovat Power BI Mobile -sovellus sekä Power BI -palvelu, joka tukee raporttien yhteismuokkaamista ja jakamista useammalle henkilölle. (Microsoft 2021a.)

Ohjelmistoon voidaan yhdistää erilaisia tietolähteitä (Excel, PDF, tekstitiedosto/CSV, Oracle/SQL tietokanta jne.). Data ladataan ohjelmistoon tai käytetään Direct Query toimintoa, jolla Power BI on suoraan yhteydessä datalähteeseen ja kysely tehdään datalähteestä. Suurien tietomäärien kohdalla Direct Query -toiminnon käyttäminen voi olla huomattavan hidasta eikä toiminto ole tuettu kaikkien tietolähteiden kohdalla. (Microsoft 2021a.)

Power BI Query Editorissa käyttäjä rakentaa erilaisia kyselyitä ja muuntaa tietoja ennen raporttien muodostamista. Datan muokkausominaisuudet ovat ohjelmassa kuitenkin rajalliset ja suuren datamäärän muokkaus on hidasta tai datan määrästä riippuen jopa mahdotonta. Raportit muodostetaan Power BI Desktop ohjelmassa, jonka jälkeen niitä voi jakaa joko Power BI Servicen kautta tai erilaisina tiedostona. (Microsoft 2021a.)

Power BI:n hinnoittelussa on erilaisia vaihtoehtoja. Power BI Desktop työpöytäsovellus on ilmainen ja sen lisäksi Power BI:llä on kolme eri hinnoittelutasoa: ilmainen Power BI Free, Power BI Pro sekä joko käyttäjäkohtaisesti laskutettava Power BI Premium Per User tai käytetyn kapasiteetin mukaan laskutettava versio. (Microsoft 2021b.)

Ilmaisen version avulla voi käyttää Power BI ohjelmistoa ja sen kaikkia analytiikkaominaisuuksia. Power BI -palvelussa käyttäjä voi luoda omia raportteja ja koontinäyttöjä. Ilmainen versio on tarkoitettu henkilökohtaiseen käyttöön. Kaikki jakamisominaisuudet ovat poissa käytöstä, pois lukien mahdollisuus julkaista julkisia hallintapaneeleja verkossa. (Microsoft 2021b.)

Power BI Pro versio on tarkoitettu pienille ja keskisuurille yrityksille. Pro -lisenssillä saa käyttöön Power BI:n kaikki ominaisuudet, mukaan lukien raporttien julkaiseminen, jakaminen ja ad hoc -analyysit. Raportteja pystyy jakamaan organisaation sisällä kaikille, joilla on Power BI Pro lisenssi. Power BI Pro:n hinta on 8,40 € käyttäjää kohden kuukaudessa. (Microsoft 2021b.)

Power BI Premium on suunniteltu suurille yrityksille. Ohjelmistolla luotua sisältöä voivat katella vain yrityksen hankkiman lisenssin piiriin kuuluvat henkilöt. Premium lisenssejä on käyttäjäkohtaisia (16,90 € käyttäjä/kuukausi) sekä käytetyn kapasiteetin mukaan laskutettavia. Kapasiteettia hankitaan nodeina. Yhden noden hinta on tällä hetkellä 4200 €/kk. (Microsoft 2021b.)

4 Kehittämismenetelmät

Laadullisen tutkimuksen tavoitteena ei ole kerätä numeerista tai määrällistä tietoa tutkittavasta kohteesta. Laadullisen tutkimuksen menetelmillä kerätty tieto on kokemuksia, mielipiteitä ja erilaisia näkökulmia tutkittavasta aiheesta. Aineistoa voidaan kerätä lisää myös aineiston analyysivaiheessa, jos huomataan jonkin näkökulman muuttuvan tai herättävän tutkijan mielenkiinnon. (Juuti & Puusa 2020, luku 2.)

Tietoa yrityksen tämänhetkisestä raportointitavasta sekä tavoitteista kerättiin henkilökunnalle tehtävillä teemahaastatteluilla. Haastattelut tehtiin yrityksen tiloissa ja jokaiseen haastatteluun varattiin aikaa yksi tunti. Haastattelujen runko hahmoteltiin etukäteen ja haastattelutilanteessa syntyvälle keskustelulle varattiin reilusti aikaa.

Teemahaastattelu valittiin menetelmäksi, koska yrityksen henkilökunnalla ei ole kattavaa tietoa Moliksen käytön mahdollisuuksista eikä rajoitteista ja siten raportoinnin mahdollisuuksista. Haastattelutilanteessa on mahdollista esittää lisäkysymyksiä sekä keskustella laajemmin raportoinnin mahdollisuuksista. Vilkkun (2015, luku 5) mukaan tutkimushaastatteluun osallistuvat henkilöt eivät ole vain tiedon keräämistä varten, vaan myös haastateltavien tulisi saada haastattelussa uutta tietoa ja pystyä kehittämään ajattelu- ja toimintatapojaan.

4.1 Tutkimuksen luotettavuuden arviointi

Tutkimuksen luotettavuuden arvioinnin keskeisimmät termit ovat reliabiliteetti sekä validiteetti. Reliabiliteetti tarkoittaa tutkimuksen luotettavuutta. Tutkimuksen luotettavuuteen vaikuttaa se, kuinka huolellisesti ja toistettavasti valittu tutkimusmenetelmä mittaa tutkittavaa ilmiötä tai asiaa. Validiteetti käsitteenä tarkoittaa sitä, mitataanko tutkimuksessa sitä mitä on tarkoitus mitata. Validiteettiin vaikuttaa esimerkiksi valittu tutkimusmenetelmä. (Hirsjärvi, Remes & Sajavaara 2008, 226-227; Tilastokeskus 2021a; Tilastokeskus 2021b.)

Tutkimuksen luotettavuuden arvioimisessa kiinnitetään huomiota tutkimuksen raportointiin. Tutkimusraportissa tulee olla perusteltuna käytetyt menetelmät ja tutkimuksen eteneminen tulee olla kuvattuna kattavasti. Lukijan täytyy ymmärtää tutkijan tekemät ratkaisut sekä saada luotettava ja selkeä kuva tutkimuksen toteutuksesta ja tuloksista. (Juuti & Puusa 2020 luku 2).

4.2 Eettisyys

Henkilötietojen käsittelystä opinnäytetyössä on omat määräyksensä ja säädöksensä. Henkilötietoja ovat sellaiset tiedot, joiden perusteella henkilö voidaan tunnistaa joko suoraan tai välillisesti. Sellaisia ovat esimerkiksi nimi, syntymäaika, henkilötunnus ja osoite. Henkilötietojen käsittelystä tulee olla tietosuojasetuksessa tai tietosuojalaissa mainittu peruste ja vain tutkimukselle välttämättömiä tietoja tulee käsitellä. Myös henkilötietoryhmien käsittelystä

tulee olla tarkkana ja niiden käsittely on lähtökohtaisesti kiellettyä. (Ammattikorkeakoulujen opinnäytetöiden eettiset suositukset. 2020, 7-11.)

Tässä opinnäytetyössä ei käsitellä henkilötietoja, mutta ne ovat työn tekijän saatavilla. Työn tekeminen vaatii pääsyä monenlaisiin potilastietojärjestelmiin. Lisäksi opinnäytetyötä tehdessä käsitellään yritysten nimiä ja niiden tilaamien tutkimusten määriä ja tutkimusnimikkeitä. Yritysten ja henkilöiden tiedot piilotetaan opinnäytetyöhön tulevista kuvista ja tilastoista. Tietoa haetaan haastattelemalla yrityksen henkilökuntaa, mutta heitä ei varsinaisesti tutkita, joten työlle ei tarvitse tehdä eettistä ennakoarviointia. Kehittämistyöhön osallistuvat tuotannon ja markkinoinnin henkilökunta, johdon henkilöstöä sekä it-osastolla työskentelevät henkilöt. (Ammattikorkeakoulujen opinnäytetöiden eettiset suositukset. 2020, 18-21; Juuti & Puusa 2020, luku 2.)

4.3 Teemahaastattelu

Teemahaastattelusta ts. puolistrukturoitu haastattelu on yleisesti käytetty tapa kerätä tutkittavaa aineistoa. Teemahaastattelussa tutkija kerää tutkimuksensa kannalta tärkeitä teemat ja aiheet, joita käsitellään haastateltavan kanssa. Aiheiden käsittelyjärjestyksellä ei ole väliä vaan tärkeintä on, että haastateltava pystyy antamaan oman näkemyksensä kustakin teemasta ja se saadaan tallennettua asianmukaisesti myöhempää analyysiä varten. Teemahaastattelu voidaan tehdä yksilö-, pari- tai ryhmähaastatteluna. Teemahaastattelua ja erityisesti ryhmäteemahaastattelua suositellaan erityisesti työelämän kehitysprojekteihin sen toimintaa kehittävän luonteen takia. (Vilkkä 2021, luku 5.)

Parihaastattelu on menetelmänä ryhmähaastattelun alalaji. Kun haastatellaan useampaa ihmistä samaan aikaan, tavoitteena on usein synnyttää keskustelua ja spontaani kommentointi tuottaa monipuolista tietoa tutkittavasta kohteesta. Tällöin tutkijan tehtävänä on enemmänkin keskustelun aikaansaaminen toisin kuin yksilöhaastattelussa tutkijan rooli on esittää kysymyksiä. Pari- ja ryhmähaastatteluilla voidaan saada aikaan yhteinen näkemys tietyistä asiasta haastateltavien kesken. (Hirsjärvi & Hurme 2005, luku 7.)

Haastattelun tekemiseen tulee valmistautua huolella. Erityisesti asiantuntijan haastattelemisessa tutkijan valmistautuminen voi vaikuttaa lopputulokseen ja haastattelun etenemiseen. Haastateltava muokkaa puhe- ja vastaustapaansa sen mukaan mikä vaikutelman hän saa tutkijasta. Tietämättömän oloiselle tutkijalle puhutaan kansanomaisin termein, kun taas asiantuntevalle puhutaan ammattitermein. Puhetapa saattaa vaikuttaa merkittävästi saatavan aineiston laatuun. (Pfadenhauer 2009, 85.)

Haastattelun alkuhetket ovat tärkeitä, koska ensivaikutelma voi vaikuttaa luotettavan ilmapiirin syntyyn ja siten vaikuttaa koko haastattelun etenemiseen. Haastattelu kannattaakin aloittaa epävirallisella keskustelulla ja sen jälkeen edetä tutkimuksen ja haastattelun tarkoituksen

kertaamiseen. Haastattelijan kannattaa myös varautua kertomaan jonkin verran itsestään ja suhteestaan tutkittavaan aiheeseen. Tutkijan avoimuus heijastuu usein myös tutkittavan puhetapaan. (Hyvärinen 2017, 39.)

4.4 Haastattelujen analyysi

Haastattelulla kerätyn aineiston määrä on usein suuri, vaikka haastateltavien määrä olisikin vain 10-15 henkeä. Usein kaikkea kerättyä materiaalia ei ole aiheellista analysoida. Analyysitapa kannattaa valita jo ennen aineiston keräämistä, jotta se voidaan ottaa huomioon haastattelua tehdessä ja näin helpottaa tulevaa tulosten analysointia. Aineiston purkaminen ja tarkastelu on suositeltavaa aloittaa mahdollisimman pian sen keruun jälkeen. Mahdolliset aineiston täydennykset ja muokkaukset on helpointa tehdä silloin kun haastattelutilanne on vielä tutkijan hyvässä muistissa. (Hirsjärvi & Hurme 2008, luku 7.)

Jos tutkija itse tekee haastattelut, analyysin voi aloittaa jo aineistoa kerätessä. Niin kutsutussa itseään korjaavassa aineiston analyysitavassa tutkija tiivistää ja tulkitsee haastateltavan kuvausta haastattelun edetessä ja kertoo tulkinnastaan myös haastateltavalle. Dialogin avulla haastattelija ja haastateltava pääsevät yhteisymmärrykseen tulkinnasta. (Hirsjärvi & Hurme 2008, luku 7.)

Aineiston analyysi voidaan jakaa kolmeen eri vaiheeseen: kuvailu, luokittelu ja yhdistely. Kuvailuvaiheessa aineistosta pyritään kuvaamaan tutkittavien kohteiden erilaisia piirteitä. Analyysin edetessä kuvattuja piirteitä jaetaan erilaisiin luokkiin teemoittain. Aineistosta etsitään yhteyksiä ja samankaltaisuuksia. Luokittelua seuraa niiden yhdisteleminen toisiinsa. Tutkija arvioi ovatko kaikki luokat tarpeellisia, löytyykö niistä yhteisiä piirteitä ja mitkä niiden keskinäiset riippuvuussuhteet ovat. Yhdistelyvaiheessa luokkien välille pyritään löytämään säännönmukaisuutta ja samankaltaisuutta. (Hirsjärvi & Hurme 2008, luku 7.)

Laadullisessa tutkimuksessa tulosten tulkintaa ja analysointia tehdään koko prosessin ajan. Tulkinta esitetään useimmiten kirjallisena raporttina. Tukena käytetään usein myös suoria lainauksia haastatteluista, kuvia, numeroita sekä kuvioita. (Hirsjärvi & Hurme 2008, luku 7.)

5 Case: BI raportointi Vita Laboratoriot Oy:ssä

Yrityksen vaatimukset raportoinnille selvitettiin keräämällä käyttäjäkokemuksellista tietoa tilastojen tarpeesta ja niiden käyttötavoista markkinoinnin ja tuotannon henkilökuntalta sekä keskustelemalla johdon kanssa. Haastateltavaksi valittiin henkilöt, jotka käyttävät työssään eniten laboratoriojärjestelmistä saatavia tilastoja. Yrityksen johdosta haastateltiin yksi henkilö ja markkinointiosastolta sekä tuotannosta kaksi henkilöä.

Tietoa nykyisistä laboratoriojärjestelmistä etsittiin käyttöohjeista, maahantuojien sekä laitevalmistajien dokumentaatiosta sekä kirjallisuudesta ja yrityksen omista dokumenteista. Osaa järjestelmien tiedoista ei ole dokumentoitu ja ne tiedot kerättiin haastattelemalla IT-osaston henkilökuntaa.

5.1 Henkilökunnan haastattelut

Henkilökunnan haastattelut toteutettiin Vita Laboratoriot Oy:n tiloissa kesän 2021 aikana teemahaastatteluina. Tuotannon sekä markkinoinnin henkilökuntaa haastateltiin parihaastatteluna, jotta samalla voitiin käydä keskustelua ja hahmotella yhteistä raportointitapaa. Johdon haastattelu toteutettiin yksilöhaastatteluna yrityksen tiloissa. Haastattelukysymykset oli hahmoteltu etukäteen erityisesti tiettyjen tarvittavien tarkkojen tietojen osalta. Haastattelut tallennettiin kirjoittamalla muistiinpanoja haastattelun aikana.

Haastattelujen aikana keskusteltiin raportoinnin mahdollisuuksista Power BI:n avulla. Haastattaville kerrottiin, minkälaista tietoa Molis -laboratoriojärjestelmästä on mahdollista hakea, sekä miten sitä voidaan käsitellä ja esittää visuaalisilla raporteilla. Haastattelun aikana nousseet raportointitarpeet kerättiin haastattelun lopuksi taulukkoon.

Kun kaikki haastatteluaineisto oli kerätty, raportointi tarpeet yhdistettiin yhteen taulukkoon ja selvitettiin, miten haluttu tieto saadaan Moliksesta. Moliksen tietokantarakennetta tutkittiin haastattelemalla Vita Laboratoriot Oy:n IT-osaston henkilökuntaa, tutustumalla ohjelman toiminnallisuuksiin sekä tekemällä tietokantahakuja Access -ohjelmiston avulla. Lisäksi selvitettiin minkälaiset visuaaliset raportointitavat ovat parhaat kullekin raportille Power BI:ssä.

Tuotannon tarpeita raportoinnin suhteen kartoitettiin haastattelemalla tuotannosta vastaavaa kemistiä sekä lajittelun tiimivastaavaa. Haastatteluissa nousi esiin tarve tiheimmälle raportoinnille sekä tyytymättömyys raportoinnin tämänhetkiseen tilanteeseen. Tällä hetkellä raporteja pyydetään muutaman kerran kuukaudessa IT-osastolta. Raporttien saamisessa on viiveitä ja saadun tiedon muokkaamiseen kuluu paljon aikaa.

Tuotannon esimiehet arvioivat tilastojen perusteella laitekapasiteetin riittävyttä, reagensien kulutusta, keskimääräisiä vastausaikoja sekä henkilökunnan riittävyttä. IT-osaston tekemän SQL-tietokantakyselyn avulla haetaan myös tiettyjen tutkimusten numeeriset vastaukset ja tehdään yleisen tulostason vertailua. Järjestelmistä saatavan datan avulla voidaan myös varautua ennakkoon tuleviin tutkimusmäärien lisääntymiseen ja vastaavasti siirtää resursseja muuhun toimintaan tutkimusmäärien laskiessa. Tutkimusmäärien tilastoja voidaan myös hyödyntää työvuorolistojen suunnittelussa.

Tuotannon esimiesten haastattelussa nousi esiin tarve tilastoida näytepoikkeamia. Väärä näytemateriaali tai tutkimuspyyntö, vanhentunut tai hyytynyt näyte vastataan järjestelmään

vakiovastauksella ("Ei tehty"), joita ei tällä hetkellä raportoida. "Ei tehty" -vastausten tilastoiminen ei ole tällä hetkellä kannattavaa, koska vastauksen syiden erittelemisen ei ole mahdollista.

Tilastojen saatavuus on tällä hetkellä huono ja riippuen IT-osaston työmäärästä, Excel muotoisen tilaston saaminen saattaa kestää useita päiviä. Tuotannon esimiehet toivoisivat seuraavia raportteja:

- Laitekohtaiset tilastot, kuinka monta näytettä on analysoitu. Tilasto jaoteltu tutkimusten sekä asiakkaiden perusteella.
- Tiettyjen tutkimusten numeeriset vastaukset jaoteltuna asiakkaittain.
- Tutkimusten vastausviiveet (tutkimus/asiakas)

Vita laboratoriot Oy:n näytteiden lajittelun esimies käyttää raportteja työmäärän arvioimiseen. Toiveena olisi päiväkohtainen raportti, josta näkyisi tunnin tarkkuudella järjestelmään kirjattujen näytteiden lukumäärä. Lisäksi vuodenaikavaihtelu näytteiden määrässä on suurta ja raporttien avulla esimerkiksi lomien suunnittelussa voitaisiin hyödyntää edellisvuosien raportteja. Lajittelun esimies toivoisi saataville raportteja, joista kävisi ilmi:

- Sisään järjestelmään kirjattujen näytteiden (putkien) määrä tunnin tarkkuudella.
- Määrän vertailu edelliseen viikkoon/kuukauteen/vuoteen

Yrityksen johto käyttää saamiaan raportteja tulevien resurssointien suunnittelemiseen. Tilastoista seurataan esimerkiksi alihankintatutkimusten määrää ja tiedon perusteella kotiutetaan tarvittaessa tutkimuksia. Lisäksi vastatut tutkimusmäärät vaikuttavat henkilöstöresurssien suunnitteluun. Tutkimusmääriä raportoidaan myös asiakasyrityksille. Raporteista seurataan myös tutkimusten vastausviiveitä.

Asiakasyritysten kanssa solmituissa sopimuksissa on määritelty tietyt vastausviiveet tutkimuksille. Johto seuraa myös näiden vastausviiveiden toteutumista. Yrityksen johto vastaanottaa tällä hetkellä raportit Excel-tiedostoina, joita muokataan tarpeisiin sopivaksi Excel Pivot taulukoihin. Johto tarvitsee seuraavia raportteja:

- Tutkimusryhmäkohtaiset tilastot/asiakas
- Tutkimuskohtaiset vastausajat ja määrät
- Vastausviive työpäivinä/tutkimus/asiakas
- Lukujen vertailu edelliseen kuukauteen/kvartaaliin/vuoteen.

Markkinointiosasto pyytää tällä hetkellä tilastoja IT-osastolta noin kerran kuussa. Tilastotietojen saamisessa on suuri viive, koska osa tutkimustilastoista saadaan laskutusohjelmasta vasta laskutuksen vahvistuttua. Asiakkaita laskutetaan kerran kuussa. (Ylisaari 2021.) Saadut Excel -

tiedostot muokataan Excel Pivot taulukoihin. Saatua dataa käytetään hintaneuvotteluissa ja sen avulla tutkitaan asiakkaiden ostokäyttäytymistä. Hintoja sekä tutkimusmääriä raportoidaan myös asiakasyrityksille. Tilastotietojen perusteella arvioidaan markkinoinnin kannattavuutta ja arvioidaan hintoja. Markkinointihenkilökunnan toiveena olisi viikoittain tai päivittäin päivittyvä raportointijärjestelmä.

- Toimialakohtaiset tutkimusmäärätilastot (esimerkiksi keskussairaalat/yksityiset/julki-
set asiakkaat)
- Kaikki saman asiakasyrityksen alayksiköiden vastatut tutkimukset
- Tutkimusryhmäkohtaiset raportit (myyntiluvut euroina ja kappalemäärinä)
- Lukujen vertailu edellisen vuoden samaan aikajaksoon sekä edelliseen viikkoon, kuu-
kauteen sekä kvartaaliin.

Haastatteluissa esiin nousseet raportointitarpeet on koottu taulukkoon 1.

Raportti:	Ketkä hyödyntävät:	Sisältö:
1. Tutkimusmäärät tutkimusryhmittäin		
Tutkimusryhmä/asiakas	Markkinointi	Tutkimukset jaoteltuina ryhmiin/asiakkaittain/kaikki vastatut tutkimukset.
2. Tutkimusmäärät asiakkaittain		
Vastatut tutkimukset/asiakas Kokonaismyynti/asiakas	Markkinointi, johto	Asiakkaittain (kaikki toimipisteet yhdistetty) jokainen tutkimus erikseen.
3. Tutkimusten vastausviive työpäivinä		
Vastausviive/tutkimus, mukaan laskettu vain työpäivät	Tuotanto, johto	Tutkimuksittain vastausviiveen keskiarvo, mediaani, keskiarvo, suurin ja pienin arvo.
4. Analysoijan kokonaistutkimusmäärä		
Analysoija/tutkimus	Tuotanto	Tutkimusmäärä per laite, yksittäisen tutkimuksen vastattu määrä.
5. Putkimäärä		
Paketit/yksittäiset pyynnöt	Tuotanto (lajittelu)	Kuinka monta putkea kirjattu sisään. Tunnin tarkkuudella. Päivä/vko/kk. Vertailu edelliseen sekä edellisvuoden samaan.
6. Hinnat		
Asiakas/tutkimus	Markkinointi, johto	Asiakkaiden toimipisteet yhdistetty.
7. Numeeriset vastaukset		
Tutkimus/asiakas	Tuotanto	Tiettyjen tutkimusten numeeriset vastaukset jaoteltuna asiakkaittain.
8. Toimialakohtaiset tilastot		
Vastatut tutkimukset/toimiala	Johto/markkinointi	Vastatut tutkimukset jaoteltuina toimialakohtaisesti: Julkinen/yksityinen

Taulukko 1: Power BI:llä toteutettavat raportit

5.2 Raportoinnin lähtötilanne

Datahakujen tekeminen laboratoriojärjestelmä Moliksesta on tällä hetkellä Vita Laboratoriot Oy:n IT-osastolla pääasiassa yhden ihmisen vastuulla. Moliksen tietokanta on Oraclen relaatio-tietokanta, jonka tarkkaa rakennekuvausta ei ole saatavilla. Tärkeimpien taulujen rakenne on tiedossa, mutta sitä ei ole dokumentoitu. IT osastolla on vakiintuneita SQL hakuja, joilla tietoa Moliksesta haetaan. (Pekkola 2021.)

Datahaut tehdään ODBC-yhteydellä Access -ohjelmiston avulla SQL datahakuna. Muodostunut Excel-tiedosto toimitetaan sitä pyytäneelle henkilölle sähköpostilla tai jakamalla verkkokansion kautta. Markkinointiosaston, johdon ja tuotannon henkilökunta muokkaa Excel-tilukon itse tarpeisiinsa sopiviksi. Datahakuja tehdään tällä hetkellä muutamia kertoja kuukaudessa. (Pekkola 2021.)

5.3 Tiedonhaku Moliksesta

Haastattelujen perusteella muodostettavia raportteja varten Moliksesta tarvitaan taulukossa 2 luetellut attribuutit.

Taulu tietokannassa	Attribuutti	Sisältö	Raporttimalli
MOLIS_ANA			
	MC	Tutkimusnimikkeen moliskoodi	1-8
	SHORTNAME	Tutkimusnimikkeen lyhenne	1-8
	EXSS	Analysaattoritieto (puuttuu osalta tutkimusnimikkeitä)	4
MOLIS_COR			
	CORNB	asiakasyrityksen koodi	1-4, 5-8
	CORN_ST	asiakkuuden aktiivisuus status	1-4, 5-8
	CONTRACT	hinnasto, jota asiakasyrityksen kaikki alayksiköt käyttävät	1-4, 5-8
	NAME1	Asiakasyrityksen nimi	1-4, 5-8
MOLIS_ORDANAEL			
	RECEPT_DT	Näytteen sisään kirjaus pvm.	3, 5
	RECEPT_HH	Näytteen sisään kirjaus klo	5
	RPT_DT	Näytteen vastaus pvm.	1-8
	RPT_HH	Näytteen vastaus klo	3
	RES_NUM	tutkimuksen numeerinen vastaus	7
	SAMPLE_ID	näyttenumero	5, 7

Taulukko 2: Moliksesta haettavat attribuutit

Taulusta MOLIS_COR haetaan seuraavat tiedot: NAME1 -kentästä saadaan koko yrityksen nimi. CONTRACT - hinnasto, jota kaikki yrityksen alayksiköt käyttävät. Sen avulla voidaan niputtaa alayksiköt yhteen ja näin saada koko yrityksen tilasto. CORN_ST on status, joka on =1 jos asiakkuus on aktiivinen ja 0 jos asiakas ei enää lähetä näytteitä ja se on inaktivoitu.

Tutkimusmäärät saadaan hakemalla Moliksesta vastatut tutkimukset asiakasyrityksittäin. Yksiköt saadaan yhdistettyä koko asiakasyrityksen yhteisiksi luvuiksi hakemalla asiakkaan käytettävää hinnastoa (CONTRACT).

Toimialakohtaisia lukuja ei voida hakea, koska järjestelmään ei ole määritelty sille attribuuttia. Toimialoittain jaetun tilaston saaminen vaatii tiedon päivittämistä Molis ohjelmaan tai jaon tekemistä ETL-prosessissa. Toimialakohtaiseen jaotteluun voidaan käyttää ohjelmassa esimerkiksi kenttää COR_SEL1.

Myös tutkimusryhmiin jakaminen olisi hyvä tehdä jo operatiivisessa järjestelmässä ennen ETL-prosessin suorittamista. Tutkimusryhmälle ei ole ohjelmassa määriteltyä kenttää. Yksi mahdollisuus tehdä tutkimusryhmäjako olisi jako analysaattoritiedon perusteella. Joillakin analysaattoreilla tehdään kuitenkin esimerkiksi sekä mikrobiologiaan, että allergiatutkimusten ryhmään kuuluvia tutkimuksia. Jakotapaa päätettäessä tulee ottaa huomioon ajankäyttö ja työhön saatavilla olevat resurssit. Moliksessa on käytössä noin 1600 tutkimusnimikettä ja niiden manuaalinen päivittäminen on työlästä.

Raporttimalli 4 (analysaattorikohtaiset tilastot) vaatii toteutuakseen sen, että Molikseen päivitetään jokaiselle tutkimusnimikkeelle analysaattoritieto. Tieto sijaitsee attribuutissa EXSS ja tietueen nimi laboratoriojärjestelmässä on "Executing subsystem". Useat tutkimukset kulkevat näytteiden lajitteluautomaatin kautta, jolloin niille ei ole määritelty EXSS tietoa. Näille tutkimuksille voidaan perustaa Molikseen näennäinen analysaattori, jonka avulla tutkimukset saadaan jaettua raporteissa oikein.

Toimiala- ja analysaattorikohtaiset jaottelut voidaan määritellä myös ETL-prosessissa. Tällöin jaottelu eri ryhmiin tehdään tietovarastojärjestelmässä. Sen ylläpitäminen ja päivittäminen uusien tutkimusten kohdalla on kuitenkin työläämpää kuin Molis-ohjelmassa tapahtuva kategorisointi.

5.4 Datan laadun muokkaus ja ETL-prosessi

Datan laadun muokkaus tulisi tehdä mahdollisuuksien mukaan jo operatiivisessa järjestelmässä. Uusien asiakasyritysten sekä tutkimusnimikkeiden syöttämisestä järjestelmään tulee tehdä yksiselitteiset ohjeet, jotta kaikki tilastointiin tarvittava tieto saadaan järjestelmään oikein ja ETL-prosessin virheidenmäärä saadaan mahdollisimman pieneksi. (Khan ym. 2012, 678.)

Ennen ETL-prosessin aloitusta laboratoriojärjestelmästä tulee inaktivoida vanhat asiakkuudet sekä käytöstä poistetut tutkimukset. Lisäksi jokaiselle asiakkuudelle tulee olla määriteltynä oikea hinnasto. Jokaisella tutkimusnimikkeellä tulisi olla määriteltynä analysaattori (EXSS), hinta, kokonainen tutkimuksen nimi sekä tutkimusnimike (SHORTNAME). Asiakkaantiedoissa pitää olla määriteltynä oikea nimi (Name1) ja mitä hinnastoa se käyttää (CONTRACT). Näiden tietojen puuttuessa raporteja ei voida toteuttaa. Lisäksi toimialakohtaiset tilastot vaativat uuden attribuutin määrittelyä järjestelmään. Moliksesta haetaan vain vastattujen tutkimusten tilastot.

Osa näytteistä ei voida tutkia esimerkiksi näytteen virheellisen kuljetuksen tai väärän näytemateriaalin takia. Tässä tapauksessa näytteeseen vastataan "Ei tehty" tai lisätään jokin muu tilannetta selittävä kommentti. Näyte tulostuu vastattuihin tutkimuksiin, mutta sitä ei laskuteta asiakkaalta. Näiden tapauksien määrä on kuitenkin suhteellisen pieni, eivätkä ne vääristä tilastoja merkittävästi. Jos "Ei tehty" näytteet halutaan pois tilastoista, tulee sitä varten määritellä ETL-prosessissa sääntö niiden poistamiseksi.

ETL-prosessissa täytyy tehdä datan muokkauksia euromääräisten myyntilukujen saamiseksi Moliksesta. Erityisesti muokkauksia vaativat pakettitutkimukset. Lääkäri pyytää potilaasta yksittäisiä laboratoriotutkimuksia sekä niin sanottuja pakettitutkimuksia. Paketeista osa on vaikiintuneita kuten esimerkiksi S-KIPA (kilpirauhastutkimuspaketti, joka sisältää S-T4-V sekä S-TSH). Asiakasyritys voi kuitenkin tilata näitä alatutkimuksia myös erikseen.

Pakettitutkimukset toteutetaan Molikseen siten, että asiakas tilaa otsikkotutkimuksen (S-KIPA) ja pyynnön saapuessa laboratoriojärjestelmään, alatutkimusten lisäävä ohjelma laukeaa ja ne lisätään pyyntöön. Koko paketin vastaukset lähetetään asiakkaalle otsikkotutkimuksen mukana.

Pakettitutkimuksen hinta on suhteessa halvempi kuin yhteenlaskettujen erillisten alatutkimusten hinta. ETL-prosessissa täytyy ensin hakea esimerkiksi kaikki S-KIPA, S-T4-V ja S-TSH-tutkimusten vastausmäärät. Sen jälkeen luvuista vähennetään S-KIPA-pakettien verran S-T4-V ja S-TSH-tutkimuksia, jotta tutkimuksia ei lasketa tilastoon kahteen kertaan. Yksittäisten tutkimusten sekä S-KIPA paketin hinnat haetaan asiakkaan hinnastosta ja kerrotaan lukumäärällä. Tämä muokkaus täytyy tehdä myös muiden pakettitutkimusten kohdalla, jotta hinnat saadaan muodostumaan raportteihin oikein.

Prosessiin määriteltyjä sääntöjä on kuvattu taulukossa 3. ETL-prosessissa tulee määritellä säännöt, joissa määritellään ensisijainen avain, attribuuttien nimet sekä mitä tapahtuu jos kenttään on syötetty useita eri arvoja. ETL-prosessissa tulee myös määritellä miten pakettitutkimusten, "Ei tehty" vastattujen, nolla ja ns. null arvojen kanssa menetellään. Lisäksi eri attribuuttien tietotyypin on oltava yhdenmukaiset. (Khan, Ehsan, Mirza & Sarwan 2012, 678.)

Määrittely	Esimerkki
Ensisijainen avain	Näytetunnus tai tutkimuskoodi puuttuu
Attribuuttien nimien määrittely	First name = FN vai Etunimi=EN?
Kenttään syötetty useita arvoja	”Koulukuja 10, 04400 Järvenpää” Postinumero ja kaupunki tulisi olla eri kentissä.
Kirjoitusvirheet	Eriksson vai Erikson?
Tyhjät arvot	Null arvo jossain kentässä
Tietotyyppien tulee olla yhteneväiset	Esim: näytenumero: numeraalinen (Int) vai teksti (String) -tyyppinen?
Päivämäärämuoto	pp/kk/vv vai pp/kk/vvv vai kk/pp/vvvv?
Sukupuoli -attribuutti	0/1 vai F/M?
Pakettitutkimukset	Yksittäisten tutkimusten määrästä vähennetään paketti otsikoiden määrä.
Vastatut, mutta tutkimattomat näytteet	”Ei tehty” -vastatut näytteet poistetaan tilastosta

Taulukko 3: ETL prosessissa määriteltävät säännöt

5.5 Tietovarastoratkaisu

Paikan päällä (engl. On-Premises) toimiva tietokanta vaatii alussa suhteellisen suuria investointeja sekä tietotaitoa järjestelmän rakentamiseen ja ylläpitoon. Se on kuitenkin usein toiminnaltaan nopeampi kuin pilvessä toimiva tietovarasto. Toisaalta pilvessä sijaitsevan tietovaraston skaalautuvuus on parempi sekä sen kustannukset jakautuvat tasaisemmin käyttöajalle. Tietovarastoratkaisujen luotettavuuteen vaikuttaa se, minkälaista osaamista yrityksen sisällä jo on ja kuinka paljon projektiin halutaan panostaa. Koulutus sekä laitteiden hankinta maksavat, mutta vaikuttavat merkittävästi projektin onnistumiseen. Tietovaraston toiminta on samanlaista molemmissa malleissa. (Hovi ym. 2009, 130, 138-139.)

Miltei kaikki markkinoilla olevat tietovarastointiratkaisut tukevat Oraclen SQL tietokantaan yhdistämistä. Tietovarastointityökalujen valintaan kannattaa käyttää aikaa ja suosia

ratkaisua, joka sopii yhteen aiempien käytettyjen ohjelmistojen kanssa. Suurimpia ETL-työkalujen tarjoajia ovat esimerkiksi IBM InfoSphere DataStage, Informatica Power Center, Microsoft SSIS ja Oracle Data Integrator. Lisäksi markkinoilla on avoimen lähdekoodin ohjelmistoja kuten Apache Kafka, Apache NiFi, CloverETL, Jaspersoft, Pentaho Kettle ja Talend Open Studio. (Päivinen 2021.)

Tietovaraston valintaan vaikuttaa moni seikka kuten millaisia tietoja varastoon tallennetaan, kuinka paljon dataa tallennetaan, kuinka nopeasti tietoja tarvitaan, kuinka paljon tietovarastoon pystytään panostamaan rahallisesti sekä ajallisesti. Monet luetelluista tekijöistä vaikuttavat suoraan toisiinsa. Esimerkiksi pienemmän skaalautuvuuden valitseminen voi heikentää suorituskykyä, mutta on yleensä kustannustehokkaampaa. Laboratorion tuottama tieto on strukturoitua tietoa, joka on tehokasta tallentaa tähtimällin relaatiotietokantaan. (Hovi ym. 2009, 36.)

Keskitetty yksi tietovarastoratkaisu sopii ainakin toistaiseksi toimeksiantajayrityksen tarpeisiin. Kerran vuorokaudessa, yön aikana päivittyvä tietovarasto riittää käyttäjien tarpeisiin eikä kerralla ladattavan datan määrä kasva silloin liian suureksi (Enterprise Data Warehouse). (Hovi ym. 2009, 27; Päivinen 2021.)

5.6 Power BI raportit

Käyttäjille jaettavat raportit toteutetaan Power BI ohjelmistolla. Datan tulisi olla valmiiksi siivottu ja muokattu tietovaraston ETL-prosessissa ennen Power BI:hin siirtämistä. Raportit muokataan ja toteutetaan Power BI Desktop ohjelmalla. Ohjelmalla muodostetaan yhteys tietovarastoon ja tietojen päivitys ajastetaan tapahtuvaksi joka yö. Tiedot päivittyvät Power BI:hin kuitenkin siten, että tietovaraston datahaku Moliksesta sekä ETL prosessi on ensin suoritettu loppuun.

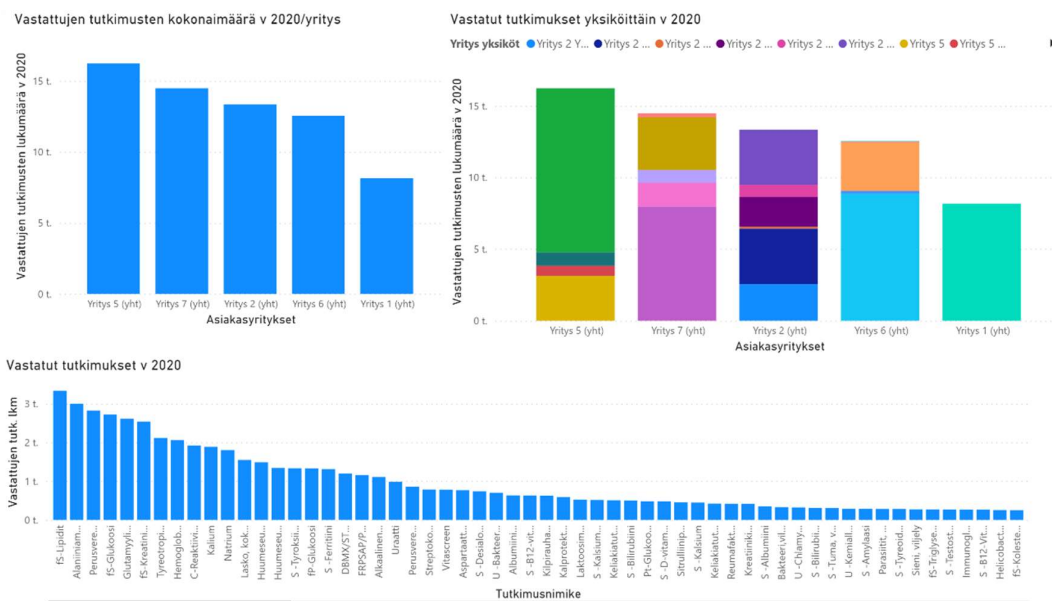
Haastatteluissa esiin nousseiden seikkojen perusteella muodostetaan kahdeksan erilaista raporttia toimeksiantajayrityksen tarpeisiin. Raportit julkaistaan verkossa Power BI -palvelussa. Raportteihin pääsy voidaan määrittellä organisaatiossa hyvinkin yksityiskohtaisesti. Kaikki raportit ovat johdon, esimiesten ja markkinointiosaston saatavilla. Koko muulle henkilökunnalle jaetaan raportit 1-5 esimerkiksi PDF -tiedostona tai Power Point -esityksenä tai ne voidaan käydä läpi palaverissa.

Seuraavat malliraportit on muodostettu Power BI ohjelmistolla Moliksesta haetun datan pohjalta. Yritysten nimet sekä tutkimusten hinnat ja määrät on muutettu tunnistamisen välttämiseksi.

Ensimmäisessä raportissa kuvataan asiakasyritykselle vastattujen tutkimusten määrä tutkimusryhmittäin. Tutkimusnimikkeet niputetaan ryhmiin erillisen ETL-prosessissa määritellyn

taulukon avulla, ohjelmallisesti Moliksessa tai jako tehdään analysointitiedon perusteella. Ryhmäjako voisi olla esimerkiksi seuraava hematologia, kemia, mikrobiologia, koronatutkimukset, allergiatutkimukset sekä alihankintana tehtävät tutkimukset (alihankintapaikoittain jaoteltuina). Vastattujen tutkimusten määrää pystyy vertaamaan edellisen kuun, sekä edellisen vuoden tilastoihin.

Raporttimalli 2 (kuvio 3) esittää kaikki vastatut tutkimukset asiakasyrityksittäin jaoteltuna. Käyttäjä voi tarkastella tietoja eri tavoin jaoteltuna aktivoimalla osioita pylväsdiagrammeista.

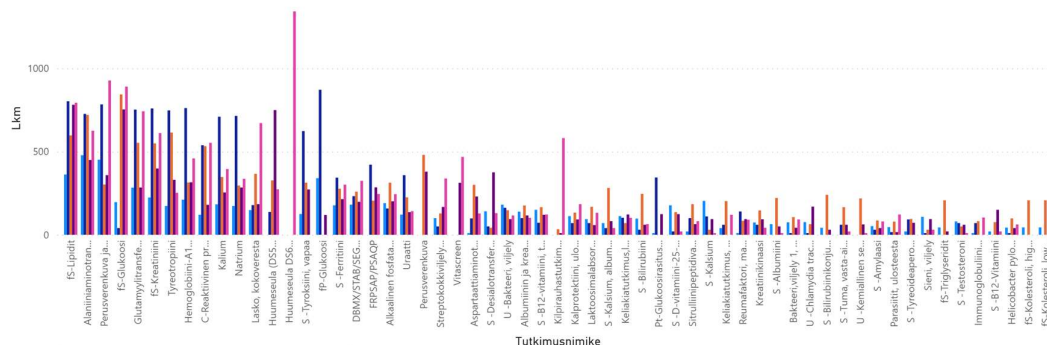


Kuvio 3: Raporttimalli 2, vastatut tutkimukset/asiakasyritys

Raporttimallin toisella sivulla kuvataan kuukausittaiset tutkimusmäärät (kuvio 4). Hiirellä aktivoimalla käyttäjä voi korostaa tietyn tutkimuksen sekä kuvaajasta, että taulukosta.

Vastatut tutkimukset/asiakasryitys/tutkimusnimike v 2020

Yritys yhteensä ● Yritys 1 (yht) ● Yritys 2 (yht) ● Yritys 5 (yht) ● Yritys 6 (yht) ● Yritys 7 (yht)

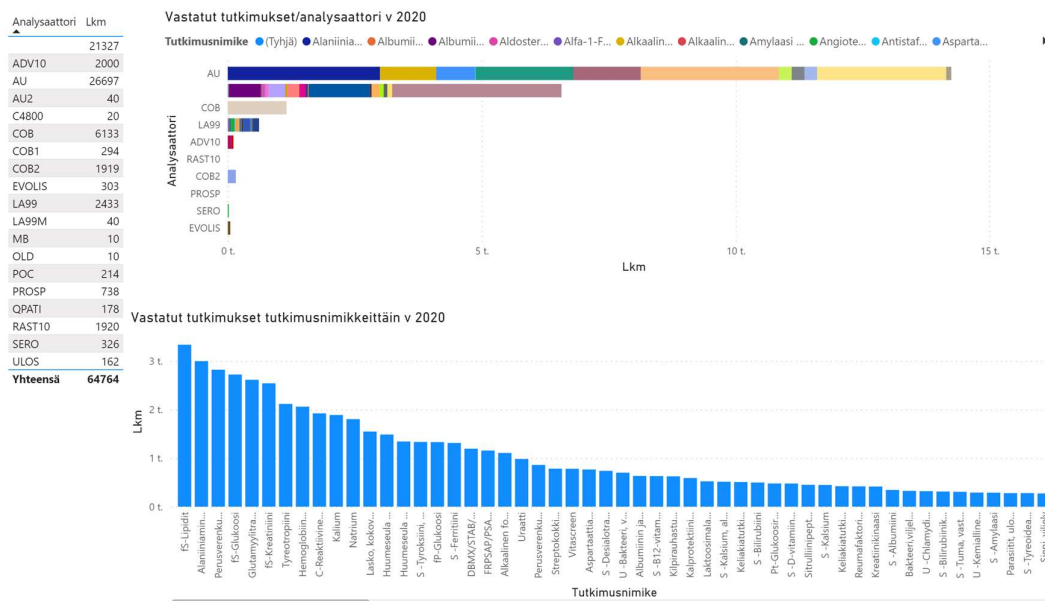


Yritys yhteensä	Tutkimusnimike	Lkm
Yritys 1 (yht)	Alaniiniaminotransferaasi	478
Yritys 1 (yht)	Albumiiniin ja kreatiniiniin suhde	139
Yritys 1 (yht)	Alkaalinen fosfataasi	190
Yritys 1 (yht)	Aspartaattiamiinotransferaasi	10
Yritys 1 (yht)	B - Fosfatidyylietanololi	49
Yritys 1 (yht)	Bakteeri, viljely 1, ulosteesta	75
Yritys 1 (yht)	BORRGE/BORRME/BORRGG/BORRMB	10
Yritys 1 (yht)	C - Reaktiivisen proteiiniin, nlaemasta	10
Yhteensä		64764

Kuvio 4: Raporttimalli 2, Vastatut tutkimukset

Raporttimallissa 3 kuvataan tutkimusten vastausviive. Laskennasta vähennetään viikonloput sekä muut pyhäpäivät, jolloin saadaan todellisten työpäivien määrä. Jokaiselle tutkimukselle on asetettu tavoitteellinen vastausaika. Vastausviive saadaan vähentämällä vastauspäivämäärästä sisään kirjauksen päivämäärä. Jos erotus on suurempi kuin asetettu tavoiteaika, luvusta tulee positiivinen. Negatiivinen luku kuvaa tavoiteaikaa nopeammin vastattua tutkimusta.

Raporttimalli neljässä (kuvio 5) kuvataan vastattujen tutkimusten kokonaismäärä jaoteltuna analysaattoreittain. Verkkopalvelussa jaetussa raportissa käyttäjä voi aktivoida hiirellä esimerkiksi tietyn tutkimuksen yläreunan palkkikuvaajasta, jolloin muissa sivun kuvaajissa näkyy kyseisen tutkimuksen tiedot korostettuina.



Kuvio 5: Raporttimalli 4, tutkimusmäärät/analyysaattori

Raporttimalli 5 esittää lajittelussa laboratoriojärjestelmään kirjattujen näytteiden määrän tunnin tarkkuudella. Tilastoa voi verrata edelliseen viikkoon, kuukauteen sekä vuoteen. Raporttiin tulostetaan myös viivakaavio koko edeltävän vuoden sisällä kirjatuista näytteistä.

Raporttimallissa 6 (kuvio 6) nähdään asiakasyrityskohtaiset hinnastot. Raportista voi selata tietyn yrityksen koko tutkimusvalikoiman hinnastoa tai listata tutkimusnimikkeen kaikki hinnat. Esimerkkikuvassa kaikkien tutkimusten hinnaksi on asetettu 1,1e.

Yritys yhteensä	Hinnat	Lyhyt
Yritys 2 (yht)	1,1	B -HbA1c
Yritys 5 (yht)	1,1	B -HbA1c
Yritys 6 (yht)	1,1	B -HbA1c
Yritys 7 (yht)	1,1	B -HbA1c
Yritys 1 (yht)	1,1	B -La
Yritys 2 (yht)	1,1	B -La
Yritys 5 (yht)	1,1	B -La
Yritys 6 (yht)	1,1	B -La
Yritys 7 (yht)	1,1	B -La
Yritys 1 (yht)	1,1	B -Lakt-D
Yritys 2 (yht)	1,1	B -Lakt-D
Yritys 5 (yht)	1,1	B -Lakt-D
Yritys 6 (yht)	1,1	B -Lakt-D
Yritys 7 (yht)	1,1	B -Lakt-D
Yritys 1 (yht)	1,1	B -PEth
Yritys 5 (yht)	1,1	B -PEth

Kuvio 6: Raporttimalli 6, hinnastot

Raporttimalli 7 sisältää Vita laboratoriot Oy:n asiantuntijoiden valitsemien tutkimusten numeristen vastausten tilastot. Vastauksista esitetään vuoden, kuukauden sekä viikon ajanjaksoilta lasketut keskiarvo, mediaani sekä suurin ja pienin arvo. Luvut esitetään kaikkien yritysten ja alayksiköiden osalta myös erikseen. Lukuja voi verrata muiden yritysten sekä eri ajanjaksojen tuloksiin.

Raporttimallissa 8 esitetään vastatut tutkimukset jaoteltuina toimialakohtaisesti. Tämä raportti vaatii toteutuakseen toimialan määrittämistä asiakasyrityksille joko ETL-prosessissa tai Molis-ohjelmassa.

6 Yhteenveto

Opinnäytetyön tavoitteena oli selvittää, minkälaisia raportteja Vita Laboratoriot Oy:n henkilökunta tarvitsee laboratoriojärjestelmästä, minkälaisia vaatimuksia käytetyt järjestelmät asettavat tietovarastolle ja mitä tietoja Moliksesta täytyy hakea, jotta tarvittavat raportit saadaan muodostettua.

Tietoa tarvittavista raporteista kerättiin haastatteluiden sekä alan kirjallisuuden avulla. Haastattelujen toteutus oli onnistunut ja haastattelujen aikana syntynyt keskustelu oli hedelmällistä ja kehittävää. Haastatteluilla saatiin selville tarvittavat tiedot ja henkilökunta oli yhteistyöhaluista ja motivoitunutta projektin edistämisen suhteen. Parihaastattelussa syntyvän keskustelun avulla voidaan kehittää työyhteisön työtapoja ja prosesseja ja siinäkin mielestä teema-haastattelu sopi hyvin tämän tutkimuksen toteuttamiseen. Haastattelun etuihin kuuluu myös se, että haastateltavien ei tarvitse itse kirjoittaa vastauksiaan kaavakkeelle, jolloin vastauksista saadaan pidempiä sekä kattavampia.

Haastattelutilanteissa haastateltaville pystyttiin myös kertomaan projektista ja sen tavoitteista tarkemmin sekä keskustelemaan järjestelmän rajoitteista ja mahdollisuuksista. Haastatteluissa nousi esiin henkilökunnan tyytymättömyys tämänhetkiseen raportointijärjestelmään sekä tarve laajemmalle ja helpommin saavutettavalle raportoinnille. Henkilökunta oli hyvin perillä mihin ja miten saatua dataa voidaan hyödyntää päätösten tekemisessä sekä työn kehittämisessä. Tässä opinnäytetyössä kartoitetun raportointitarpeen, Moliksen tietokantarakenteen ja ETL-prosessin selvityksen avulla voidaan Vita Laboratoriot Oy:ssä edetä kaupallisten ratkaisujen kartoitukseen ja tietovaraston rakentamiseen.

7 Jatkokehittämisehdotukset

Tämä opinnäytetyö on osa toimeksiantajan projektia, jonka tavoitteena on saada kaikki tuotantodata yhteen raportointijärjestelmään. Perustettavaan tietovarastoon on mahdollista liittää myös käytössä oleva laskutusjärjestelmä. Tietovarastossa voidaan muokata dataa sellaiseen muotoon, että laskutusjärjestelmässä muodostettavat laskut palvelevat paremmin sekä asiakkaiden, että Vita laboratoriot Oy:n tarpeita. Tällä hetkellä laskutukseen liittyy paljon manuaalista tietojen syöttämistä sekä tarkistamista (Ylisaari 2021). Esimerkiksi laskujen poti-laserittelyt voisi lähettää sähköisesti asiakkaille tietovaraston kautta.

Jokaisen järjestelmän liittäminen tietovarastoon on suuri ja paljon henkilöstöresursseja vaativa projekti ja se tulee valmistella huolellisesti. Ympäristölaboratoriossa käytettävä Fision Lims -laboratorio-ohjelmisto on Microsoftin SQL tietokanta ja se on yhdistettävissä useimpiin tietovarastoratkaisuihin. Pohdittavaksi jää, kuinka suuri hyöty saadaan yhteyden rakentamisesta.

Patologian osastolla käytettävän ohjelmiston, QPatin yhdistettävyyteen ei ehditty tämän opinnäytetyön yhteydessä perehtyä. Todennäköisesti sen yhdistäminen samaan tietovarastoon olisi vuorossa viimeisenä sen vaatiman suuren työmäärän sekä taloudellisen panostuksen takia.

Tuotannon esimiesten haastatteluissa nousi esiin tarve tilastoida erilaisia näytepoikkeamia. Tällä hetkellä tilastoiminen ei ole automatisoidusti mahdollista, koska kaikki erilaiset näytepoikkeamat kirjataan Molis -laboratoriojärjestelmään samanlaisella vakiovastauksella ("Ei tehty"). Jatkokehitysprojektina Molikseen tulisi tehdä erilaiset vakiovastaukset erilaisille näytepoikkeamille, jolloin niiden tilastoiminen Power BI -ohjelmistoon on mahdollista.

Operatiivisten järjestelmien tuottaman tiedon tilastoiminen ja analysoiminen on jatkuvasti kehittyvä prosessi, joka ei valmistu koskaan. Pienilläkin laboratoriojärjestelmän ja henkilöstön työskentelytapojen muutoksilla voi olla merkittäviä vaikutuksia tietovarastointiin ja edelleen tiedon tilastointiin. Tiedon raportointiin onkin suhtauduttava jatkuvana kehitysprojektina ja siihen on myös varattava aikaa ja resursseja, jotta raportointijärjestelmä pysyy ajan tasaisena ja luotettavana.

Lähteet

Painetut

Hirsjärvi, S., Remes, P. & Sajavaara, P. 2008. Tutki ja kirjoita. Helsinki: Kirjayhtymä.

Hovi, A., Hervonen, H. & Koistinen, H. 2009. Tietovarastot ja business intelligence. Helsinki: WSOYpro.

Hovi, A., Ylinen, J. & Koistinen, H. 2001. Tietovarastot liiketoiminnan tukena. Helsinki: Satku.

Hyvärinen, M. Haastattelun maailma. Teoksessa Hyvärinen, M, Nikander, P., Ruusuvuori (toim.), J., Aho, A. L. & Granfelt, R. 2017. Tutkimushaastattelun käsikirja. Tampere: Vastapaino.

Pfadenhauer, M. 2009. At the eye level: The expert interview - a talk between expert and quasi-expert. Teoksessa Bogner, A., Littig, B & Menz, W. (toim.) Interviewing experts. Basingtoke: Palgrave Macmillan.

Törmänen, A. 1999. Tietovarastointi: Strategiasta toteutukseen. Helsinki: Suomen atk-kustannus.

Törmänen, A. 2017. Johdanto tietovarastointiin. North Charleston: CreateSpace.

Sähköiset

Ammattikorkeakoulujen opinnäytetöiden eettiset suositukset. Ammattikorkeakoulujen rehtorineuvosto Arene ry. 9.1.2020. Viitattu 10.7.2021. <https://www.arene.fi/wp-content/uploads/Raportit/2020/AMMATTIKORKEAKOULU-JEN%20OPINN%C3%84YTET%C3%96IDEN%20EETTISET%20SUOSITUKSET%202020.pdf>

CGM - CompuGroup Medical 2021. Mehr Sicherheit durch Ihr Labor. Viitattu 1.8.2021. https://www.cgm.com/deu_de/produkte/labor/molis.html

George, J., Kumar, V & Kumar, S. 2015. Data Warehouse Design Considerations for a Healthcare Business Intelligence System. Proceedings of the World congress on engineering 2015, 1-3 July, 2015. http://www.iaeng.org/publication/WCE2015/WCE2015_pp308-311.pdf

Hirsjärvi, S. & Hurme, H. 2008. Tutkimushaastattelu: Teemahaastattelun teoria ja käytäntö. E-kirja. Helsinki: Gaudeamus Helsinki University Press.

Juuti, P. & Puusa, A. 2020. Laadullisen tutkimuksen näkökulmat ja menetelmät. E-kirja. Gaudeamus.

Kimball, R., Ross, M., Becker, B., Mundy, J. & Thornthwaite, R. 2008. The Data Warehouse Lifecycle Toolkit. E-kirja. John Wiley & Sons, Incorporated.

Khan, A., Ehsan, N., Mirza, E. & Sarwan, S. Z. Integration between Customer Relationship Management (CRM) and Data Warehousing. *Procedia Technology* 1. 2012. 239-249. Viitattu 1.7.2021 <https://www.sciencedirect.com/science/article/pii/S2212017312000515/pdf>

Microsoft 2021a. Get started With Power BI Desktop. Viitattu 21.7.2021. <https://docs.microsoft.com/en-us/power-bi/fundamentals/desktop-getting-started>

Microsoft 2021b. Power BI Pricing. Viitattu 1.8.2021. <https://powerbi.microsoft.com/en-us/pricing/>

Miller, G. C., Dolan, C. T., Crowson A. N. & Stout, S. R. Data warehousing and information management strategies in the clinical immunology laboratory. *Clinical and Applied Immunology Reviews* 3. 2002. 73-86. Viitattu 2.7.2021. <https://www.sciencedirect.com/science/article/abs/pii/S1529104902000594>

Sherman, R. 2014. Business Intelligence Guidebook. E-kirja. Morgan Kaufmann.

Silvers, F. 2011. Data Warehouse Designs. E-kirja. Auerbach Publishers, Incorporated.

Souibgui, M., Atigui, F., Zammali, S., Cherfi, S & Yahia, S. B. Data quality in ETL process: A preliminary study. *Procedia Computer Science*. Volume 159. 2019. 676-687. Viitattu 13.7.2021. <https://www.sciencedirect.com/science/article/pii/S1877050919314097/pdf>

Tilastokeskus 2021a. Käsitteet - reliabiliteetti. Viitattu 1.7.2021. <https://www.stat.fi/meta/kas/reliabiliteetti.html>

Tilastokeskus 2021b. Käsitteet - validiteetti. Viitattu 1.7.2021. <https://www.stat.fi/meta/kas/validiteetti.html>

Vilkka, H. 2021. Tutki ja kehitä. E-kirja. Jyväskylä: PS-kustannus

Vita Laboratoriot Oy 2020. Tilinpäätös. Viitattu 12.6.2021.

Väre, T. 2019. Master data. E-kirja. Helsinki: Alma Talent.

Julkaisemattomat

Hurme, H. 2021. Liiketoimintajohtajan haastattelu 9.6.2021. Vita Laboratoriot Oy. Helsinki.

Pekkola, M. 2021. IT-päällikön haastattelu 12.4.2021. Vita Laboratoriot Oy. Helsinki.

Päivinen, M. 2021. ICT-päällikön haastattelu 18.6.2021. Vita Laboratoriot Oy. Helsinki.

Ylisaari, P. 2021. Talousasiantuntijan haastattelu 1.7.2021. Vita Laboratoriot Oy. Helsinki.

Kuviot

Kuvio 1: Tietovarastointi ja BI raportointiprosessi.....	10
Kuvio 2: ETL-prosessin eteneminen	12
Kuvio 3: Raporttimalli 2, vastatut tutkimukset/asiakasyritys.....	28
Kuvio 4: Raporttimalli 2, Vastatut tutkimukset	29
Kuvio 5: Raporttimalli 4, tutkimusmäärät/analysointori	30
Kuvio 6: Raporttimalli 6, hinnastot	30

Taulukot

Taulukko 1: Power BI:llä toteutettavat raportit	21
Taulukko 2: Moliksesta haettavat attribuutit.....	23
Taulukko 3: ETL prosessissa määriteltävät säännöt	26