

Henkilöstöraportoinnin kehittäminen ja automatisointi

Case: Rovaniemen koulutuskuntayhtymä REDU

LAB-ammattikorkeakoulu

Insinööri (ylempi AMK)

2022

Tapio Alhonsuo

Tiivistelmä

Tekijä(t) Alhonsuo, Tapio	Julkaisun laji Opinnäytetyö, YAMK	Valmistumisaika 2022
	Sivumäärä 57	
Työn nimi Henkilöstöraportoinnin kehittäminen ja automatisointi Case Rovaniemen koulutuskuntayhtymä REDU		
Tutkinto ja koulutusala Insinööri (ylempi AMK), Digitaaliset ratkaisut		
Toimeksiantajan nimi Rovaniemen koulutuskuntayhtymä REDU		
Tiivistelmä Työelämän muutokset luovat haasteita henkilöstön hyvinvoinnille. Työelämän alati muuttuvat vaatimukset edellyttävät työntekijöiltä ja organisaatioilta myös jatkuvaa opimista. Organisaation toiminnasta syntyvän tiedon laaja hyödyntäminen mahdollistaa resurssien kohdentamisen tavalla, joka vie organisaatiota sen toivomaan suuntaan. Opinnäytetyön tavoitteena oli tuottaa REDUn esihenkilöille kaksi spesifiä henkilöstöhallinnon raportoinnin työkalua tiedolla johtamista varten. Työkalut eli Power BI -raportit koskivat henkilöstön sairauspoissaoloja sekä henkilöstön osaamisen kehittämiseen varatun resurssin käyttämistä. Nämä halutut henkilöstöresurssin tiedot olivat olemassa REDUn tietojärjestelmissä, mutta ne eivät olleet ajantasaisesti esihenkilöiden saavutettavissa. Tavoitteista johdettiin seuraavat kaksi tutkimuskysymystä: Miten REDUn johto voisi seurata sairauspoissaoloja niin toimialatasolla kuin työntekijäkohtaisestikin, ja miten REDUn johto voisi seurata osaamisen kehittämiseen varatun resurssin käyttämistä? Tutkimuksellinen kehittämistyö toteutettiin konstruktivisella tutkimusotteella. Sen artefaktina valmistuivat tavoitteen mukaiset raportit tietosuojan vaikutustenarvioineen, mutta raportteja ei otettu tuotantokäyttöön puutteellisen tietokantayhteyden vuoksi. Tavoitteen osittainen epäonnistuminen korostaa raportointia ja analysointia varten perustettavan tietovaraston merkitystä tulevaisuuden henkilöstöraportoinnin tarpeisiin vastattaessa.		
Asiasanat henkilöstöhallinto, henkilöstöraportointi, tiedolla johtaminen, tietosuoja, datan visualisointi		

Abstract

Author(s) Alhonsuo, Tapio	Type of Publication Thesis, UAS Number of Pages 57	Published 2022
Title of Publication Development of personnel reporting and automation Case Rovaniemen koulutuskuntayhtymä REDU		
Degree and field of study Master of Engineering, Digital Solutions		
Name of the client organization Rovaniemi Municipal Federation of Education REDU		
Abstract The ever-changing demands of work challenge the well-being of employees and requires constant learning from both employees and organizations. Extensive utilization of the information generated by an organization's operations allow resources to be allocated in a way that takes the organization in the direction it desires. The aim of this thesis was to provide REDU's management with two specific personnel management tools. The tools, ie the Power BI reports, concerned the sick leave of the personnel and the usage of the resource set aside for the development of the personnel's competence. The desired human resource information existed in REDU's information systems, but was not easily available for the management in an up-to-date state. The following two research questions were derived from the objectives: How could REDU's management monitor sick leave at both the department and employee level, and how could REDU's management monitor the usage of the resource set aside for competence development? The research and development work was carried out with a constructive research approach. As an artifact, the targeted reports with data protection impact assessments were completed, but the reports were not put into production due to a lack of database connection. The partial failure of the target highlights the importance of a data warehouse for reporting and analysis in meeting future personnel reporting needs.		
Keywords human resources, data privacy, data visualization		

Sisällys

1	Johdanto.....	1
1.1	Tausta ja lähtökohdat	1
1.2	Tavoite, tutkimuskysymykset ja työn rajaus	3
1.3	Tutkimusmenetelmät	4
2	Nykytila ja haasteet.....	7
3	Tietoperusta.....	9
3.1	Tiedolla johtaminen.....	9
3.2	Henkilöstöpääoma ja henkilöstöraportointi.....	10
3.3	Datan kerääminen ja mallintaminen	11
3.4	Data-analytiikka ja Business Intelligence	14
3.4.1	Microsoft Power BI.....	16
3.5	Datan visualisointi.....	17
3.6	Tietosuoja.....	19
3.6.1	Mitä on tietosuoja?	19
3.6.2	Tietosuojalainsäädäntö.....	21
3.6.3	Rekisteröidyn oikeudet	21
3.6.4	Tietosuoja ja tietoturva.....	22
3.6.5	Henkilötietojen käsittelyn perusteet ja oikeutetun edun arviointi.....	23
3.6.6	Vaikutustenarviointi ja riskianalyysi.....	24
3.6.7	Tasapainotesti	24
3.6.8	Henkilötieto ja anonymisointi	24
3.6.9	Erityiset henkilötietoryhmät ja arkaluonteiset tiedot.....	25
3.6.10	Henkilötietojen käsittely	25
3.6.11	Henkilörekisteri	25
3.6.12	Suppeat ja laajat pysyvät henkilötietorekisterit.....	26
3.6.13	Käsittelytoimintojen luettelo	26
3.6.14	Tietosuojaselosteet.....	26
3.6.15	Osoitusvelvollisuus	27
4	Kehittämisen prosessin toteutus	28
4.1	Tutkimuksellinen kehittäminen prosessina	28
4.2	Tietosuojan toteutus	29
4.2.1	Riskianalyysin ja vaikutustenarvioinnin kehitys työn aikana	30
4.2.2	Tasapainotesti	32
4.2.3	Aineiston tietoturvallinen käsittely kehitysvaiheessa	35

4.2.4	Aineiston anonymisointi ja normalisointi.....	36
4.3	Iteraatiot anonymisoiduilla tiedoilla	37
4.3.1	Ensimmäinen iteraatio	37
4.3.2	Toinen iteraatio.....	43
4.3.3	Kolmas ja viimeinen iteraatio	45
4.4	Kehitystyön päätös	45
5	Yhteenveto ja johtopäätökset.....	47
5.1	Johtopäätökset	47
5.2	Jatkokehittäminen.....	49
5.3	Pohdinta	50
5.4	Kehittämistyön eettisyys ja luotettavuus.....	50
5.5	Oman oppimisen ja ammatillisen kehittymisen tarkastelut	52
	Lähteet	53

1 Johdanto

1.1 Tausta ja lähtökohdat

Työelämän muutokset luovat haasteita henkilöstön hyvinvoinnille. Toiminnan nopeutuminen, tuottavuuden parantaminen ja liiketoiminnan tehokkuuden maksimointi aiheuttavat paineita niin työntekijöille kuin heidän esihenkilöilleenkin (Viitala 2009, 212). 1970-luvulla teollisuus oli työn rasittavuutta koskevien tilastojen kärjessä. Nykyään johtoon ovat kirineet kunnalliset alat, kuten opetus- ja terveydenhuoltoala. Trendi on vastaava muissakin teollisuusmaissa. Tyypillisesti työhyvinvointia uhkaavat työn jatkuvuuden epävarmuus, tietotyön tuomat raskaudet, lisääntynyt kiire sekä työn ja vapaa-ajan rajan hämärtyminen. Työnantajalle nämä näkyvät poissaolojen lisääntymisenä, henkilöstön vaihtuvuutena sekä alentuneena työtyytyväisyytenä (Viitala 2009, 223-225). Poissaoloista yksistään aiheutuu työnantajalle merkittävä määrä suoria kustannuksia. Työhyvinvointia tukemalla voidaan siis säästää selvää rahaa, minkä lisäksi hyvinvoiva henkilöstö on avainasemassa laadukkaiden palveluiden tuottamisessa kuntalaisille (Viitala 2009, 213; Julkaisut Helsinki 2020).

Työelämän alati muuttuvat vaatimukset edellyttävät työntekijöiltä ja organisaatioilta myös jatkuvaa oppimista. Hyvin suunniteltu osaamisen kehittäminen varmistaa, että organisaatio pysyy muutostuulien mukana. Kunta-alan toimintaympäristössä on useita kehityssuuntia, jotka vaikuttavat henkilöstön tarvitsemien taitojen laatuun ja määrään. Muun muassa teknologiakehitys, rakenneuudistukset ja työurien pidentäminen edellyttävät osaamistarpeiden muutosten ennakkointia ja analysointia. Osaamisen kehittäminen on suunnitelmallista osaamisen ylläpitämistä, mutta myös poistuvan osaamistarpeen korvaamista tarvittavalla osaamisella sekä kokonaan uuden osaamisen hankkimista. Jotta koko organisaation osaaminen kehittyy, tulee kehittämisen toimintatapoja ja tuloksia myös arvioida. Arvioinnin perusteella toimintaa voidaan korjata haluttuun suuntaan ja osaamisen johtamista kehittää. (Kunta- ja hyvinvointialueyönantajat KT 2017.)

Kunta-alalla on tehty suositus henkilöstön osaamisen kehittämisestä. Osaamisen kehittämisen tavoitteena on palvelutuotannon tulokellisuus sekä työelämän laadun parantaminen. Osaamisen kehittämisen tulee kytkeytyä kunnan palvelu- ja henkilöstöstrategioihin ja yksilön työuraan. Osaamisen johtamisella henkilöstön osaamisen kehittäminen yhdistetään kunnan tavoitteisiin. Henkilöstön kehittämisen tavoitteena onkin turvata tarvittava osaaminen kunnan palvelutuotantoon. (Kunta- ja hyvinvointialueyönantajat KT 2015.)

Rovaniemen koulutuskuntayhtymässä REDUssa (jatkossa pelkkä REDU) on määritelty vuonna 2021 osaamisen kehittämisen uudeksi mittaritavoitteeksi jokaiselle työntekijälle 18

palkallista tuntia vuodessa. Tämän mittaritavoitteen raportoinnin kehittäminen REDUn esihenkilöille on toinen tämän opinnäytetyön päätavoitteista.

Sairauspoissaolot ovat yksi tärkeä mittari arvioitaessa henkilöstön työkyvyn ja hyvinvoinnin tilaa organisaatiossa. Jos henkilöstöltä vaaditaan liikaa, se voi lisätä sairastavuutta tai mataltaa kynnystä jäädä sairauslomalle. Syyt sairauspoissaoloihin ovatkin moninaisia, ja työnantajalla on mahdollisuus vaikuttaa vain osaan syistä. Yllättäviin sairauksiin on vaikea ellei mahdoton reagoida ennakoivasti, mutta tutkimusten mukaan myös monet työyhteisötekijät lisäävät sairauspoissaolojen riskiä. Näitä tekijöitä ovat esimerkiksi vähäinen vaikutusmahdollisuus omaan työhön, sosiaalisen tuen puute työpaikalla, organisaatiosupistukset ja työsuhteen epävarmuus. (Viitala 2009, 213.)

Syy- ja seuraussuhteet sairauspoissaoloissa ovat harvoin selviä kuten voidaan nähdä vuotta 2020 tarkastellessa, jonka keväänä koronaviruspandemia alkoi. Yli 35 000:sta COVID-19:ään sairastuneesta huolimatta KELA:n korvaamat pitkät sairauspoissaolot vähenivät Suomessa viisi prosenttia edelliseen vuoteen verrattuna (Johns Hopkins University of Medicine 2021; YLE 2021). Syyksi on arveltu lisääntyntä etätyötä, joka on osalla työntekijöistä parantanut työhyvinvointia ja vähentänyt työtapaturmien sattumista, mutta toisaalta se on myös helpottanut sairaana töiden tekemistä mikä ei näy tilastoissa. Pandemian myötä ihmisten kontaktit vähenivät, mikä taas vaikutti esimerkiksi flunssan leviämiseen. Muutoksia tapahtui myös hoitoon hakeutumisessa ja hoitoon pääsyssä. (YLE 2021.)

Sairauspoissaolojen lisääntymisestä ei siis voi vetää suoria johtopäätöksiä ilmiön syistä, mutta se voi heijastaa jotain organisaatiossa tapahtunutta muutosta. Sairauspoissaolojen raportoinnin kehittäminen REDUn esihenkilöille onkin toinen tämän opinnäytetyön tavoitteista.

Näiden kahden edellä mainitun raportoinnin työkalun toivottiin mahdollistavan tiedolla johtamista, jossa päätöksentekoa tukisi ajantasainen tieto ja sen analysointi. Organisaation toiminnasta syntyvän tiedon laaja hyödyntäminen mahdollistaisi muun muassa resurssien kohdentamisen tavalla, joka veisi organisaatiota sen toivomaan suuntaan (Jalonen 2015, 1). Ajantasainen tieto oli olemassa REDUn tietolähteissä, mutta se ei ollut helposti kaikkien siihen oikeutettujen saavutettavissa. Ajantasainen henkilöstöraportointi mahdollistaisi esihenkilöille henkilöstövoimavaran kohdentamisen siten, että toiminnalle asetetut tavoitteet olisivat paremmin saavutettavissa (Armstrong 2006, 5).

Näiden raportoinnin työkalujen ensimmäinen tavoitetilä määriteltiin REDUn osaamisen kehittämisen työryhmän toimesta, jonka kanssa varsinaista kehitystyötä myös tehtiin.

Kehittämistyössä käytettiin konstruktivistista tutkimusotetta. Metodologiana se keskittyy reaallimaailman ongelmiin, jotka koetaan käytännössä tarpeellisiksi ratkaista.

1.2 Tavoite, tutkimuskysymykset ja työn rajaus

Opinnäytetyön tavoitteen määritteli REDUn osaamisen kehittämisen työryhmä. Kyseessä oleva työryhmä toimii asiantuntija- ja valmisteluryhmänä, pyrkien toiminnallaan kehittämään muun muassa REDUn henkilöstön osaamisen hallinnan, tasa-arvon ja työhyvinvoinnin kehittämisen yhteisiä linjauksia. Työryhmän valmistelemat asiat päätetään REDUn hallintosäännön ja sen delegointipäätösten mukaisesti. Osaamisen kehittämisen tavoitteena on niin työntekijän kuin työyhteisön osaamisen turvaaminen sekä laadun ja tuloksellisuuden varmistaminen ja parantaminen.

Työryhmä määritteli tavoitteeksi uusien työkalujen kehittämisen kahta erillistä tarkoitusta varten. Näillä työkaluilla johdon tulisi pystyä seuraamaan henkilöstön kehittämiseen tarkoitettujen resurssin toteutumista, sekä seuraamaan sairauspoissaoloja niin toimiala- kuin työntekijätasollakin. Tarkemmin avattuna, esihenkilön tulisi pystyä seuraamaan osaamisen kehittämiseen resursoitujen tuntien kertymistä juuri hänen alaisilleen. Samaten esihenkilön tulisi pystyä seuraamaan omien alaistensa sairauspoissaoloja, mutta ei muiden.

Työkalujen toivottiin mahdollistavan myös sen, että sairauspoissaolojen vuodenaikoihin tai muihin aikamääreisiin liittyvät mahdolliset trendit tulisivat ilmi. Tiettyyn yksikköön tai sijaintiin liittyvä sairauspoissaolojen keskittyminen tulisi myös pystyä havaitsemaan.

Tavoitteesta johdetut tutkimuskysymykset ovat:

- Miten REDUn johto voisi seurata osaamisen kehittämiseen varatun resurssin toteutumista?
- Miten REDUn johto voisi seurata sairauspoissaoloja niin toimialatasolla kuin työntekijäkohtaisestikin?
 - Miten REDUn johdolle mahdollistetaan työhyvinvointihankkeiden kohdentamista tukevan tiedon löytäminen?
 - Miten REDUn johdolle mahdollistetaan vuodenaikoihin tai muihin aikamääreisiin liittyvien sairauspoissaolotrendien löytäminen?

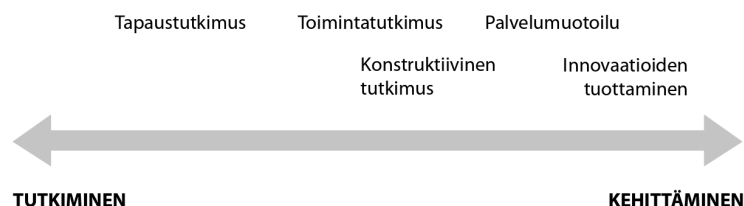
Tavoite sovittiin toteutettavaksi kehitystyönä, jonka lopputuotteena syntyisi johdon käyttöön valmiita Microsoft Power BI -raportteja. Power BI oli jo valmiiksi käytössä REDUssa eri yhteyksissä, eikä sen käyttöönotto tähän kehitystyöhön edellyttänyt hankintoja.

1.3 Tutkimusmenetelmät

Ylempien ammattikorkeakoulututkintojen opinnäytetyöt ovat tyypillisesti soveltavan tutkimuksen töitä, koska ammattikorkeakoulujen opinnot suunnitellaan työelämän tarpeiden mukaan. Soveltavassa tutkimuksessa sovelletaan perustutkimuksen tuottamaa tieteellistä tietoa johonkin käytännölliseen ongelmaan, ja sen päämääränä ovat käytännön sovellukset (Tieteen termipankki 2016). Ammattikorkeakoulun englanninkielinen nimi University of Applied Sciences tarkoittaaakin soveltavan tutkimuksen korkeakoulua. Soveltavaan tutkimukseen yhdistetään usein myös kehittämistoimintaa. Soveltavan tutkimuksen tutkimussuunnitelma on usein joustava, kuitenkin siten, että metodologinen läpinäkyvyys on jatkuvasti näkyvässä roolissa.

Tutkimuksellinen kehittämistoiminta sijoittuu työelämän kehittämistoiminnan ja tutkimuksen välimaastoon. Kun työelämän kehittämistoiminnassa sovelletaan tutkimuksissa saatua tietoa työelämän kehittämis- ja uudistamistarpeisiin, tutkimus taas tuottaa uutta tietoa ja asioita, joida voidaan soveltaa käytäntöön. Tutkimuksellisessa kehittämistoiminnassa yhdistyy konkreettinen kehittämistoiminta ja tutkimuksellinen lähestymistapa. Uusi tieto tuotetaan käytännön toimintaympäristössä, jossa apuna toimivat tutkimukselliset menetelmät. (JAMK 2021.)

Tutkimuksellisella kehittämistyöllä voidaan pyrkiä ratkaisemaan käytännöstä nousseita ongelmia tai uudistamaan ja luomaan tietoa työelämän käytännöistä. Kehittämisen tueksi kehitetään systemaattisesti ja kriittisesti arvioimalla tietoa niin käytännöstä kuin teoriastakin. Kehittämistyössä ei ainoastaan pyritä kuvailemaan ja selittämään asioita, vaan siinä etsitään niille parempia vaihtoehtoja ja viedään asioita käytännössä eteenpäin. Aktiivinen vuorovaikutus kehittämistyön eri osapuolten kanssa korostuu. Tutkimuksellinen kehitystyö eroaa tieteellisestä tutkimuksesta pääasiallisesti toiminnan päämäärissä: tavoitteena ei ole niinkään saada aikaan tieteellistä teoriaa, vaan ennen kaikkea käytännön parannuksia ja uusia ratkaisuja. Tämä ero vaikuttaa kehitystyössä käytettäviin lähestymistapoihin ja menetelmiin. Tyypillisiä kehitystyössä käytettäviä lähestymistapoja ovat muun muassa tapaustutkimus, toimintatutkimus, konstrukttiivinen tutkimus, palvelumuotoilu ja innovointi. Kehittämistehtävän luonne määrittelee käytettävän lähestymistavan. (Ojasalo ym. 2015, 18.) Siinä missä tieteellisen tutkimuksen teoreettinen viitekehys nojautuu pelkästään tutkimuskirjallisuuteen, tutkimuksellinen kehittäminen voi hyödyntää myös työelämän tuottamaa tietoa (Humanistinen ammattikorkeakoulu 2021). Erilaiset kehitystyön lähestymistavat sijoittuvat tutkimisen ja kehittämisen välisellä jatkumolla kuvion 1 mukaisesti.



Kuvio 1. Lähestymistavat tutkiminen-kehittäminen -jatkumolla (Ojasalo ym. 2015, 36)

Konstruktiiivisen tutkimusotteen ydinkäsite on konstruktio, eli jokin uusi, abstrakti käsite, jolla on ääretön määrä mahdollisia toteutumia. Kaikki ihmisen luomat artefaktit, kuten mallit, suunnitelmat, rakenteet, tuotteet tai vaikkapa tietojärjestelmämallit, ovat konstruktioita. Luonteenomaista niille on, että niitä ei löydetä, vaan ne kehitetään. Konstruktiiivisen tutkimusotteen ideaalinen tulos on, että jokin reaali maailman ongelma ratkaistaan implementoidulla uudella konstruktiolla. Tulee kuitenkin muistaa, että käytännön tason toteutuksen epäonnistuminen on yksi mahdollinen lopputulema. Kehitystyö voi silti olla akateemisesta näkökulmasta mielenkiintoinen. Tällöin tulee kuitenkin pohtia, miksi ongelmanratkaisuprosessi epäonnistui, ja tunnistaa, mitä muutoksia ennakkoehtoihin tarvittaisiin, jotta voitaisiin löytää toimiva ratkaisumalli. (Lukka 2000, 3.)

Tutkimuksessa tuotettu konstruktio on siis ratkaisu johonkin esitettyyn ongelmaan. Konstruktio käsitteenä sisältää myös toteuttamisyrityksen, jolla testataan sen käytännön soveltuvuutta. Konstruktiiivisessa tutkimusotteessa odotetaan tapahtuvan myös kokemuksellista oppimista, joka kumpuaa tutkijan ja käytännön edustajien vuorovaikutuksesta. Tutkimuksen tulee olla huolellisesti kytketty olemassa olevaan teoreettiseen tietämykseen, ja empiiriset löydökset tulee reflektoida takaisin teoriaan. (Lukka 2001.)

Toimeksiantajan REDUn toimesta tämän opinnäytetyön tavoitteeksi määriteltiin uusien työkalujen kehittäminen spesifiä tarkoitusta varten. Koska tavoite oli jotain konkreettista, jotain mikä piti kehittää koska sitä ei sellaisenaan ollut saatavilla, oli konstruktiiivinen tutkimusote luonnollinen metodologiavalinta. Kehitystyön konstruktio oli REDUn esihenkilöille heidän käyttöoikeuksiensa mukaan rajatut BI-raportit, joiden tarkempi sisältö ja tarkoitus on kuvattu tavoitteessa luvussa 1.2.

Tutkimuksellisen kehittämistyön prosessi voidaan jakaa kronologisesti kuuteen eri vaiheeseen tai osa-alueeseen. Ensimmäinen vaihe on kehittämiskohteen tunnistaminen ja alustavien tavoitteiden määrittäminen, jonka jälkeen kehittämiskohteeseen perehdytään teoriassa ja käytännössä. Tämän jälkeen kehittämistehtävä määritellään ja kehittämiskohde rajataan. Sitten on vuorossa tietoperustan laatiminen sekä lähestymistavan ja menetelmien suunnittelu, joita seuraa kehittämishankkeen toteuttaminen ja julkistaminen eri muodoissa.

Viimeisenä vaiheena on kehittämisprosessin ja lopputulosten arviointi. Mikäli kehitystyö jatkuu, prosessi aloitetaan alusta. (Ojasalo ym. 2015, 24.)

Opinnäytetyö sisälsi myös tapaustutkimuksellisia piirteitä. Tapaustutkimuksessa ei pyritä hankkimaan suurta ja mahdollisimman edustavaa aineistoa tutkittavasta asiasta, vaan paneudutaan syvällisesti rajatumpiin tapauksiin ja tutkitaan niitä monipuolisesti eri kulmilta. Sillä voidaan yrittää ymmärtää ilmiöitä, jotka ovat vahvasti sidoksissa kontekstiinsa, kuten REDUn tarve edistyneempään henkilöstöraportointiin. Ilmiön ymmärtämisessä on olennaista huomioida sen taustat ja olosuhteet, kuten esimerkiksi organisaation tietotekninen ympäristö tai kunta-alan taloudellinen ja sosiaalinen tilanne. Tapaustutkimuksella voidaan yrittää vastata kysymyksiin ”miten” ja ”miksi”. Se tuottaa tyypillisesti yksityiskohtaista tietoa hyvin rajatusta aiheesta. (Saaranen-Kauppinen & Puusniekka 2022.)

2 Nykytila ja haasteet

REDUn henkilöstötietoa on paljon, mutta se on koottuna eri järjestelmiin. Tiedosta raporttien tuottaminen vaatii paljon manuaalista työtä ja raportit ovat käyttäjien eli esihenkilöiden kannalta monen mutkan takana. Osa tiedoista kerätään vain vuosittain esimerkiksi henkilöstöraportin tai tilinpäätösseurannan yhteydessä. Joidenkin tietojen osalta tämä voi riittää, mutta joidenkin tietojen osalta ajantasaisempi tieto helposti saavutettavassa muodossa mahdollistaisi enemmän tietoon perustuvaa johtamista.

REDUn nykyinen raportointityökalujen valikoima ei mahdollista sitä, että esihenkilö voisi vapaasti haluamanaan ajankohtana saada ajantasaisen tiedon alaistensa osaamisen kehittämisen resurssin toteutumisesta tai sairauspoissaolojen статистиikasta. Ajantasaisen tiedon saavuttaminen vaatii kolmannen osapuolen eli henkilöstöpalvelujen osallistamista. Etenkin osaamisen kehittämisen kertymistä haluttaisiin seurata sitä mukaa, kun resurssia käytetään, tai ennen kaikkea ollaan käyttämättä, jolloin esihenkilöllä olisi mahdollisuus vaikuttaa tilanteeseen suotuisasti. Yksittäisten henkilöiden tietojen kyselemisen sijaan esihenkilöt hyötyisivät yleisnäkymästä, joka koskisi heidän kaikkia alaisiaan. Tällöin he voisivat paremmin muodostaa oman kokonaiskuvansa tilanteesta.

EU:n yleisen tietosuoja-asetuksen sekä kansallisen tietosuojalain nojalla REDUn kuuluu organisaationa arvioida henkilötietojen käsittelyyn liittyviä riskejä, eli tietosuojariskejä, ennen kuin se ryhtyy käsittelemään henkilötietoja uudella tai uudistetulla tavalla. Tämän opinäytetyön kehitysosassa suunniteltiin kahden uuden käsittelytoiminnon käyttöönottoa, ja näiden toimintojen tietosuojariskien arviointi tehtiin kuten minkä tahansa uuden käsittelytoiminnon vaikutustenarviointi. REDUn oman tietosuojaohjeistuksen mukaisesti tietosuojariskien arviointi on tehtävä muun muassa hankintojen yhteydessä, kun hankitaan uutta tietojärjestelmää tai -palvelua tai hankitaan henkilötietojen käsittelyä ulkopuoliselta palveluntoimittajalta, jossa käsitellään REDUn asiakkaiden, opiskelijoiden tai henkilökunnan henkilötietoja uudella tavalla tai yhdistetään aiempia henkilötietoja.

Tietosuoja vaikutusten arvioinnilla ja siihen liittyvän riskianalyysin avulla voi tunnistaa jo suunnitteluvaiheessa ne toimenpiteet, joihin on ryhdyttävä riskien hallitsemiseksi tai niiden poistamiseksi ja henkilötietojen asianmukaisen käsittelyn turvaamiseksi. Tietosuoja-asetuksen mukainen riskiarvio on tehtävä rekisteröidyn näkökulmasta eli rekisterinpitäjän on arvioitava, mitä rekisteröidyn vapauksia ja oikeuksia käsittely voi vaarantaa, ja mitä vahinkoja rekisteröidylle voi aiheutua suunnitellusta henkilötietojen käsittelystä.

REDUn oman tietosuojaohjeistuksen mukaisesti tehdyssä alkukartoituksessa toinen uusista käsittelytoiminnoista tunnistettiin vähäisen riskin toiminnoksi, mutta toinen tunnistettiin

mahdollisesti korkean riskin aiheuttavaksi toiminnoksi. Jälkimmäiselle tehtiin siten myös laajempi riskianalyysi ja tietosuojavaikutusten arviointi (TVA), joilla arvioidaan käsittelyn tarpeellisuutta, oikeasuhteisuutta ja henkilötietojen käsittelystä aiheutuvia riskejä sekä tarvittavia toimenpiteitä, joilla riskeihin puututaan. Tavoitteena oli arvioida ja arvioinnin tuloksen perusteella päättää, onko jäljelle jäänyt riski oikeutettu ja hyväksyttävissä. Vaikutustenarvioinnin tarkoitus on auttaa rekisterinpitäjää tietosuojalainsäädännön vaatimusten noudattamisessa, sen dokumentoinnissa ja osoittamisessa.

3 Tietoperusta

3.1 Tiedolla johtaminen

Tiedolla johtamiselle ei ole yksiselitteistä määritelmää. Organisaatiot saattavat kärsiä niin tiedon puutteesta kuin sen paljoudestakin (Laihonen ym. 2013, 15). Tiedolla johtamisella pyritään vähentämään epävarmuutta, joka johtuu informaation puutteesta tai sen paljoudesta ja monitulkintaisuudesta (Jalonen 2015, 1). Se on arvon luontia aineettomista voimavaroista, ja siihen kuuluu tiedon tuottaminen, hallinta, säilyttäminen, analysointi ja hyödyntäminen. Ennen kaikkea se on kyky toimia tiedon ja tietämyksen pohjalta. (Kosonen 2019, 4.)

Tiedolla johtaminen voidaan erotella tietojohdamisesta, joskin näitäkin termejä käytetään ristiin keskenään. Tietojohdaminen ry, joka pyrkii edistämään tiedon hyödyntämistä päätöksenteossa, määrittelee tiedolla johtamisen tiedon hyödyntämiseksi päätöksenteossa, kun taas tietojohdamisen arvon luonniksi aineettomista voimavaroista (Tietojohdaminen Ry. 2015). Tietojohdaminen tai tiedolla johtaminen ei ole kuitenkaan itseisarvoista. Sen oikeutus syntyy organisaation kyvystä erotella olennainen informaatio epäolennaisesta ja kyvystä hyödyntää tätä informaatiota organisaation toiminnan parantamiseen. (Jalonen 2015, 1.) Prosessin tulee siis tukea organisaatiota sen tavoitteiden saavuttamisessa. Resursseja ei tule haaskata tarkoituksettomaan tietojen keruuseen tai analysointiin. Oleellista on oleellisen tiedon hyödyntäminen. (Laihonen & Lönnqvist 2013.)

Tiedolla johtamiselle ei myöskään ole suoraa englanninkielistä vastinetta, mutta sen alle sijoittuvia kapeampia substanssialueita on. Business intelligence on liiketoimintatiedon hallintaa, joka tarjoaa keinoja organisaation toimintaympäristössä olevan informaation seurantaan ja analysointiin päätöksenteon tueksi. Knowledge management eli tietämyksenhallinta keskittyy jo olemassa olevan tiedon ja osaamisen jakamiseen organisaation sisällä. Intellectual capital management eli aineettoman pääoman johtaminen auttaa organisaation johtoa tunnistamaan organisaation arvokkaimmat aineettomat resurssit ja kehittämään niitä kohti ennalta määriteltäviä tavoitteita. Myös innovaatiojohtamisella (eng. innovation management) on yhtymäpintaa tiedolla johtamiseen, koska siinä tietoa pyritään muuttamaan uusiksi ideoiksi ja innovaatioiksi. (Wikipedia 2020.)

Tässä opinnäytetyössä kehitetyt Power BI -raportit pyrkivät mahdollistamaan tiedon hyödyntämisen esihenkilöiden päätöksenteossa. Tämä pyrittiin saavuttamaan muun muassa tarjoamalla informaatiota helpommin saavutettavassa muodossa. Jotta REDU organisaationa voisi hyötyä opinnäytetyön konstruktiosta, pyrittiin kehitystyön aikana erottelemään

epäolennainen tieto olennaisesta, ja tarjoamaan esihenkilöille ainoastaan tarkoituksenmukaista informaatiota.

3.2 Henkilöstöpääoma ja henkilöstöraportointi

Aineettomalla pääomalla tarkoitetaan sellaista yrityksen tai organisaation hallussa olevaa resurssia tai voimavaraa, jota ei huomioida tilinpäätöksessä varallisuutena. Henkilöstöön sidottua aineetonta pääomaa voidaan pitää organisaation tärkeimpänä voimavarana (Kesti 2010, 19). Yksi osa aineetonta pääomaa on niin kutsuttu inhimillinen pääoma, jolla viitataan henkilöstön osaamiseen ja kyvykkyyteen. (Wikipedia 2021.)

Yrityksen inhimillinen pääoma muodostuu työntekijöiden osaamisesta, kuten tutkinnoista ja kumuloituneesta ammattitaidosta, mutta myös heidän henkilökohtaisista ominaisuuksistaan, kuten yritteliäisyydestä ja asenteesta. Koska inhimillinen pääoma on sidottu yksittäisiin ihmisiin, joiden motiiveista tai kyvyistä organisaatio ei välttämättä ole tietoinen, ei organisaatio aina voi täysin hallita sitä, kuten se voi vaikka kiinteistöomistuksiaan. Inhimillinen pääoma voidaan kuitenkin tunnistaa, sitä voidaan pyrkiä mittaamaan, ja mittaamisen onnistuessa sen nykytilaa voidaan arvioida suhteessa tavoitetilään. (Kujansivu ym. 2007, 28, 46.) Mittaamista varten pääoman elementeille pitää kyetä antamaan tiettyä ajankohtana jonkinlainen niiden laatua tai tilaa ilmaiseva määre, ja mittaus pitäisi pystyä toistamaan, useasti, jotta mitattavan asian kehitystä olisi mahdollista ymmärtää. Monia aineettomaan pääomaan liittyviä asioita on pitkään jo mitattukin yrityksissä. Näitä ovat esimerkiksi sairauspoissaolot, henkilöstön vaihtuvuus tai työilmapiirin muutokset. (Viitala 2009, 329.) Viitala (2009, 329) esittää mittaamisen olevan kuitenkin tyypillisesti sattumanvaraista tai pelkäänsä henkilöstöhallinnon toimesta tapahtuvaa, minkä sijaan mittaamisen tulisi olla yrityksen tai organisaation strategiaan ja pitkän aikavälin tavoitteisiin sidottua tarkkaan suunniteltua toimintaa. Hyvien inhimillisen pääoman mittareiden hän kuvailee olevan sellaisia, jotka muun muassa ovat johdon ja esihenkilöiden tärkeiksi kokemia, ja jotka kertovat myös henkilöstön tuottokyvystä, eivätkä pelkäänsä kustannuksista, ja jotka ovat ymmärrettäviä ja helposti viestittäviä, ja joiden tuottaminen on helppoa ja taloudellisesti edullista.

Armstrong (2006, 70) toteaa: *"You cannot manage unless you measure"*. Vapaasti suomentuna hän tarkoittaa sitä, että henkilöstövoimavaran tunteminen on edellytys sen hallinnoinnille. Henkilöstöhallinnollisten käytäntöjen ja toimenpiteiden vaikuttavuuden arviointia pyritään seuraamaan henkilöstöraportoinnilla. Henkilöstöhallinnollisen tiedon analysointi auttaa organisaatiota löytämään, ylläpitämään, kehittämään ja kohdentamaan yhtä sen suurimmista resursseista eli henkilöstövoimavaraa. (Armstrong 2006, 70.)

Vuonna 2013 Kunta-alan työnantajajärjestö KT Kuntatyönantajat antoi suosituksen henkilöstövoimavarojen arviointiin. Kuntatyönantajan tulisi tuntea henkilöstövoimavaransa, jotka mahdollistavat kunnan tuottamat palvelut. Henkilöstövoimavarojen ennakointi on osa strategiatyötä, jolla pyritään turvaamaan työvoiman saatavuus, henkilöstön osaaminen, työhyvinvointi ja työn tuloksellisuus. Tieto henkilöstövoimavaroista auttaa henkilöstösuunnittelua palvelutarpeiden muutoksissa. (KT Kuntatyönantajat 2013.)

Henkilöstöhallinnollisia asioita mitattaessa tulisi kuitenkin jatkuvasti arvioida mittauksen tarjoama arvo organisaatiolle ja sitä, ovatko mittaamisen seuraukset niin yksiselitteisiä kuin alun perin on ajateltu. Onko itse mittaamiseen käytetty energia ja resurssi oikeassa suhteessa siitä saatavaan hyötyyn? Mitataanko jotain, mihin ei kuitenkaan ole mahdollisuuksia vaikuttaa? Kuinka paljon tiedon kerääminen vaikuttaa itse mittaustulokseen? Puutteelliseen tai vääristyneeseen tietoon perustuvat päätökset saattavat aiheuttaa organisaatiolle ylimääräisiä haasteita kehittymisen sijaan. (Banfield & Kay 2012, 245.)

Lisäksi aineettoman pääoman mieltäminen joksikin staattiseksi resurssiksi, kuten vaikka raaka-ainevarastoksi tai taitopankiksi, on ongelmallinen. Yrityksessä voidaan esimerkiksi ajatella olevan tietty määrä osaamista. Riskinä on tällöin huomion kohdistaminen aineettoman pääoman määrään ja laatuun, eikä niiden muutoksiin. Aineeton pääoma, kuten henkilöstön tutkinnot, eivät kuitenkaan itsessään tuota lisäarvoa organisaatiolle. Arvon tuottaminen on mahdollista vasta, kun tutkintojen tuomaa lisäosaamista hyödynnetään johonkin. Aineeton pääoma onkin luonteeltaan alati muuttuvaa ja kehittyvää. Esimerkiksi henkilöstön osaaminen voi kouluttautumisen kautta kehittyä, mutta se voi myös vähentyä mitään tekemättä, muun maailman kehittyessä ympärillä. (Kujansivu ym. 2007, 30.) Asioihin voidaan kuitenkin vaikuttaa. Osaamisen kehittämiseen kannustavan ilmapiirin luominen organisaatioon vaatii pitkäjänteistä ja suunnitelmallista työtä. Palkitsemis- ja kannustinjärjestelmät sekä osaamisen kehittyessä urakehityksen mahdollistaminen ovat perinteisiä keinoja, joiden lisäksi vaaditaan vahvaa sitoutumista organisaation johdolta. (Österberg 2005, 120.)

Tämän opinnäytetyön tavoitteena oli kehittää ja automatisoida henkilöstöraportointia. Tuotoksena syntyi kaksi raporttia, jotka mittasivat aineetonta pääomaa. Raportit koskivat henkilöstön sairauspoissaoloja sekä henkilöstön osaamisen kehittämiseen varatun resurssin toteutumista suhteessa sille asetettuun tavoitteeseen.

3.3 Datan kerääminen ja mallintaminen

Ennen kuin dataa voidaan tutkia, se tarvitsee kerätä. Ennen keräämistä tulisi miettiä kaikki oleelliset lähteet, joista dataa voi olla saatavilla. Riittääkö ainoastaan yksi tietolähde, vai voisiko jokin muu lähde täydentää datasta tehtäviä analyyseja? Pääsääntöisesti mitä

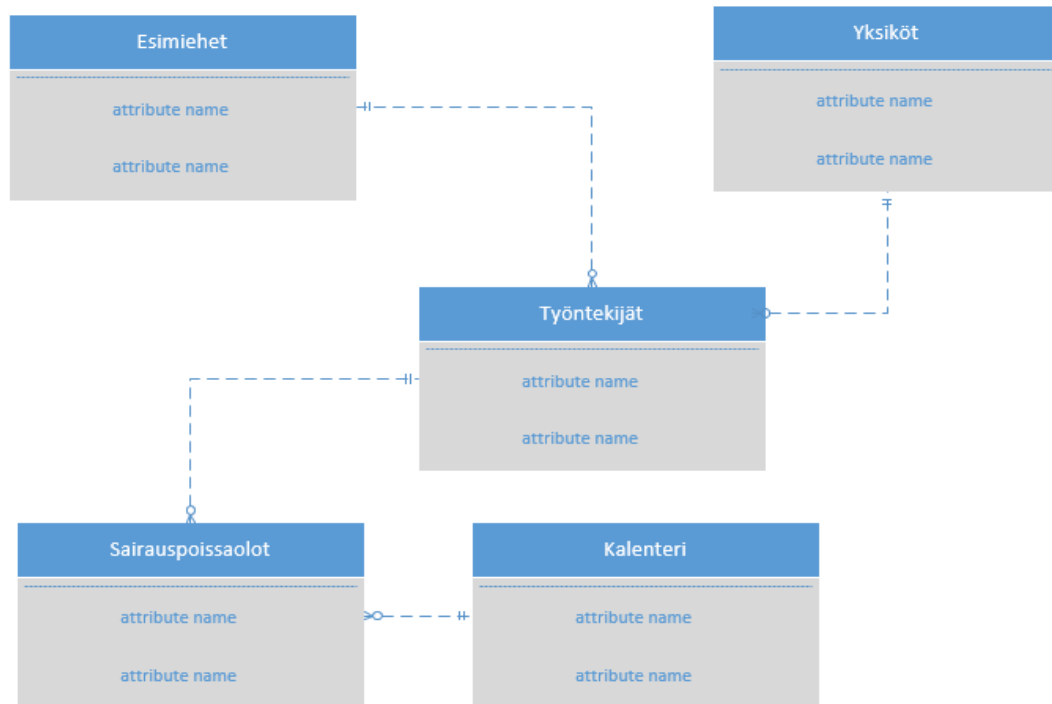
enemmän dataa on, sitä parempi. Dataa voi aina karsia ja suodattaa myöhemmin, mikäli se nähdään tarpeelliseksi.

Tosielämän data hyvin harvoin käy sellaisenaan analysoitavaksi. Siinä voi olla aukkoja, havaintojen kaksoiskappaleita, mittaus- tai kirjausvirheitä tai muita epäjohdonmukaisuuksia, jotka haittaavat tai jopa estävät jatkokäsittelyn. Englanninkielinen käsite ”GIGO” kuvaa tämän vaiheen tärkeyttä: ”Garbage in, Garbage out”, eli sotkuinen data tuottaa sotkuisia malleja ja tulkintoja. Dataa siistittäessä jokainen vaihe tulee kin perustella ja dokumentoida, jotta sen vaikutukset ymmärretään myös myöhemmissä vaiheissa. (Baesens 2014, 13.)

Datan mallintamisella tarkoitetaan käsittelyn kohteena olevan datan rakenteiden ja ominaisuuksien yhteyksien kuvaamista. Tarkoituksena on havainnollistaa minkälaista dataa on tallennettu, minkälaisia suhteita eri tyyppisillä tiedoilla on, ja miten data voidaan ryhmitellä ja organisoida sen muodon ja ominaisuuksien perusteella. Datan mallintamiseen vaikuttaa sen toivottu käyttötarkoitus, joka voi olla esimerkiksi liiketoiminnallinen tai tutkimuksellinen. Datan mallinnuksessa käytetään standardisoituja kaavioita ja muotoja, jotta mallinnuksen dokumentointi olisi johdonmukaista, yhtenevää ja ymmärrettävää.

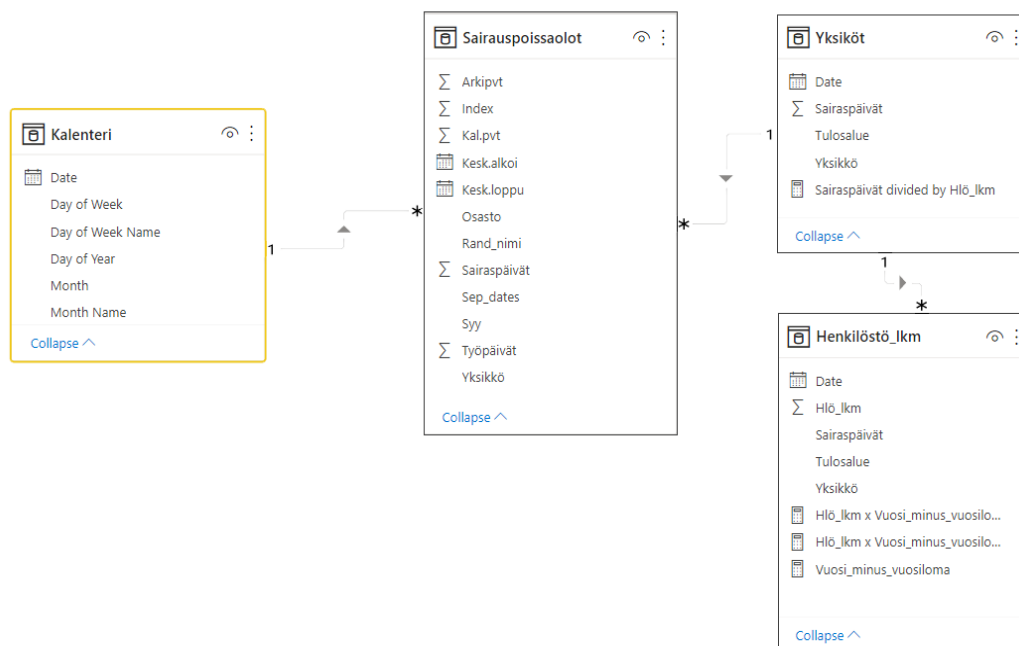
Kuten kaikissa suunnitteluprosesseissa, datan mallintamisen suunnittelu aloitetaan abstraktilta, yleiseltä tasolta, josta edetään kohti tarkempaa ja konkreettisempaa kuvausta. Datamallit voidaan karkeasti jakaa kolmeen eri kategoriaan, joilla on eri tarkkuus. Ensimmäinen on käsite- tai konseptitason malli, josta edetään rakenne- tai loogisen tason malliin, jonka kautta päädytään edelleen tallennus- eli fyysisen tason malliin.

Käsittemalli ryhmittelee datan haluttujen ominaisuuksien perusteella, ja määrittelee suhteet näiden ominaisuuksien välille (IBM Cloud Education 2020). Se voi vastata esimerkiksi kysymyksiin ”mitä tietoja käsitellään?” ja ”miten nämä tiedot liittyvät toisiinsa?”. Ryhmitelty data esitetään tyypillisesti ryhmänä eri tauluja. Suhteiden ja ryhmiteltyjen ominaisuuksien määrittelystä käytetään myös nimitystä ER-mallinnus (Entity – Relationship Modeling), ja näin määritellyistä tietovarastoista puhutaan relaatiotietokantoina. Suhteiden, eli relaatioiden kuvaamiseen BI-toteutuksissa käytetään tyypillisesti ns. tähti- tai dimensiomallia, joka tukee hyvin moniulotteista ja numeerista tietoa. Se on lisäksi selkeä ja helppo ymmärtää. (Hovi ym. 2009, 36.) Käsitetason malli ei määrittele teknisiä yksityiskohtia ja on ohjelmistoista riippumaton kuvaus tiedoista. Käsitetason malli (Kuvio 2) kuvaa mitä tietoja käsitellään ja minkälaisia suhteita tiedoilla on.



Kuvio 2. Käsitetason malli (mukailtu IBM Cloud Education 2020)

Rakennetason malliin datan erilaiset ominaisuudet määritellään tarkemmin. Malliin määritellään onko jokin tieto vaikka numeerista, tekstimuotoista tai kategorista, ja minkä ominaisuuden eli avaimen kautta suhde eli relaatio on muihin ryhmiin eli tauluihin. Tallennustason malliin (Kuvio 3) tieto kuvataan tätäkin tarkemmin.



Kuvio 3. Tallennustason malli Power BI:n esittämänä

Fyysinen- eli tallennustason malli kuvaa tarkasti miten tieto tallennetaan tietovarastoon. Se on viimeistelty esitys joka voidaan implementoida relaatiotietokannaksi. Siinä kuvataan min-käläisina koneenläheisinä rakenteina tiedot tallennetaan ja miten niitä voidaan käsitellä. (Kujala 2005.)

Tietovarastoinnilla tarkoitetaan tietojen tallentamista omaan tietokantaansa, joka on suunniteltu ja rakennettu jotain tiettyä tarkoitusta varten, kuten tässä tapauksessa tietojen raportointia varten. Tietovarasto sisältää operatiivisista järjestelmistä jalostettua tietoa, joka on tyypillisesti muokattu raportointi- ja kyselykäyttöön paremmin sopivaksi ja se päivittyy määräjain, esimerkiksi kerran päivässä. Tietovaraston tiedot ovat ainoastaan lukukäytössä. Niitä ei siis muokata suoraan, vaan ne päivittyvät ainoastaan operatiivisen järjestelmän tietojen päivittymisen mukana. Tietovaraston tietoihin kohdistetaan kyselyitä, analyysejä ja niistä voidaan koota raportteja erilaisilla BI-työkaluilla. (Hovi ym. 2009, 14.)

Tässä kehitystyössä kehitettyjen raporttien tietolähteet tuli mallintaa, ennen kuin niitä voitiin täysimääräisesti hyödyntää. Esimerkiksi yksiköiden tiedot olivat toiminnallisista syistä eri taulussa kuin yksikön työntekijöiden sairauspoissaolot. Osa tiedoista tuli taas kokonaan eri lähteistä. Näiden eri taulujen suhteet ja ominaisuudet tuli määritellä, ennen kuin niihin voitiin kohdistaa kehitystyön tavoitteiden mukaista laskentaa tai visualisointia.

3.4 Data-analytiikka ja Business Intelligence

Kuten tiedolla johtamisella, myöskään pelkällä datalla, datan keräämisellä tai sen analysoinnilla ei ole itseisarvoa. Datan arvo, oli sitä sitten paljon tai vähän, syntyy vasta asiayhteyden ja esitystavan myötä (Roberts 2012). Data on siis tulkittavissa vasta kun se muuttuu informaatioksi. Datasta luodaan informaatiota luomalla sille rakenne. Informaatio taas muuttuu edelleen tietämykseksi, kun informaatiota tulkitaan. Laihonen, ym. (2013, 18) jaottelevat nämä tiedon eri tasot kuten taulukossa 1:

TIEDON TASO	MÄÄRITELMÄ
Tietämys	Inhimillistä tietoa, joka usein perustuu kokemukseen
Informaatio	Rakenteellista dataa, jota voidaan käyttää analyyseissa
Data	Rakenteettomia tosiasioita

Taulukko 1. Tiedon eri tasot

Tietämys, jota käytetään päätöksenteossa, edellyttää siis tarkoituksenmukaisia analyysejä. Ollakseen luotettavia nämä analyysit edellyttävät tarkoituksenmukaista informaatiota.

Ollakseen tarkoituksenmukaista tämän informaation tulee pohjautua oikeanlaiseen dataan. Oikeanlainen data on relevanttia, tarkkaa ja oikea-aikaista. (Power, 2016.)

Power (2014, 225) jaottelee päätöksentekoa tukevat analyysit kolmeen eri luokkaan:

1. Takautuva (retrospective): kvantitatiivisten työkalujen käyttö menneisyydestä olevan datan ymmärtämiseen, jotta voidaan tehdä päätelmiä tulevaisuudesta.
2. Ennustava (predictive): tulevaisuuden skenaarioiden mallintaminen menneen datan perusteella. Ennustava analyysi auttaa tekemään tulevan tiedetyksi.
3. Ohjaileva (prescriptive): suunnitelmallinen kvantitatiivisten analyysien teko reaaliaikaisesta datasta joka saattaa käynnistää tapahtumia. Ohjaileva analyysi suosittelee toimia.

Organisaatioille kertyy dataa tyypillisesti useisiin eri järjestelmiin, ja sitä kertyy paljon. Eri käyttöön tarkoitettuja järjestelmiä voi olla useita: on talouden hallintaan keskittyviä järjestelmiä, opiskelijoiden opintosuoritusten hallintaan keskittyviä järjestelmiä tai vaikkapa asiakkaiden yhteystietojen ja tilausten hallintaan tarkoitettuja järjestelmiä. Useimmat näistä järjestelmistä tyypillisesti tuottavat omia raporttejaan, mutta ne eivät välttämättä onnistu tarjoamaan organisaatiolle relevanttia dataa tai ne eivät vastaa liiketoiminnan kehittämisen vaatimuksia. Kaivattu tieto voi olla vaikeasti saavutettavaa tai se ei ole käytettävissä silloin, kun sitä tarvittaisiin. (Suominen & Suominen 2015, 10.)

Business intelligencen määritelmä vaihtelee jonkin verran lähteestä riippuen. Suominen & Suominen (2015, 10) määrittelevät sen organisaation tekemäksi hallituksi datan keruuksi, ja datan muuttamista liiketoimintatiedoksi sopivilla työkaluilla. Microsoft (2022b) määrittelee sen siten, että business intelligence (BI) auttaa organisaatioita analysoimaan historiallista ja tämänhetkistä dataa, jotta sen strategiset päätökset perustuvat mahdollisimman paljon tietoon. BI mahdollistaa tämän esittämällä isoja tietojoukkoja useista eri lähteistä visuaalisessa ja helposti ymmärrettävässä muodossa. Hovi ym. (2009, 11) erottelevat BI:n tarkoitamaan ainoastaan loppukäyttäjän työkaluosuutta, ja pitävät tietovarastoinnin irrallisena koko käsitteestä. Business intelligence voi siis tarkoittaa myös koko tietojen hyödyntämisen ja analysoinnin ratkaisua, jossa tietovarastot ja niistä raportteja rakentavat toiminnot ovat vain tekninen osa BI:tä.

Microsoft jaottelee BI:n toiminnan neljään eri vaiheeseen. Näiden neljän vaiheen aikana raaka data muuttuu ymmärrettäväksi, oivalluksia tarjoavaksi tiedoksi josta jokainen organisaatioissa voi hyötyä. Ensimmäiset kolme vaihetta valmistelevat tietoa viimeistä vaihetta varten, jossa tehdään strategisia päätöksiä tiedon pohjalta. Ensimmäinen vaihe sisältää datan keräilyn halutuista lähteistä sekä sen muuntamisen yhteiseen muotoon. Tämän jälkeen

data sijoitetaan johonkin keskitettyyn paikkaan, jossa BI-ohjelmat pääsevät siihen käsiksi, ja voivat analysoida sitä yhtenä tietolähteenä. Toisessa vaiheessa pyritään löytämään datasta trendejä ja epäjohtodnmukaisuuksia. BI -ohjelmissa on tyypillisesti yksinkertaisia, automatisoituja toimintoja pintapuolisten data-analyysien tekemiseen, jotka voivat paljastaa esimerkiksi datan rakenteita tai poikkeavia havaintoja (eng. outlier). Mikäli pika-analyysit eivät tarjoa uutta tietoa tai löydettyä tietoa halutaan syventää, ohjelmista löytyy työkaluja myös esimerkiksi tilastollisten, deskriptiivisten eli kuvailevien ja prediktiiivisten, eli ennustavien, analyysien tekoon. Kolmannessa vaiheessa pyritään esittämään löydökset helposti ymmärrettävinä ja jaettavina visualisointeina, joita tyypillisesti ovat interaktiiviset raportit. Raporteissa voi olla pylväskaavioita, graafeja, karttoja tai vaikkapa tunnuslukuja, jotka auttavat käyttäjää ymmärtämään, mitä organisaation toiminnassa parhaillaan tapahtuu. Tiedot voivat päivittyä raporttiin jopa reaaliaikaisesti. Viimeisessä vaiheessa tehdään strategisia päätöksiä tiedon tarjoaman ymmärryksen myötä. Nykyhetken ja menneisyyden näkeminen oikeassa kontekstissa auttaa organisaatioita löytämään ratkaisut tämänhetkisiin haasteisiin ja mahdollisesti varautumaan tuleviin. (Microsoft, 2022b.)

Suosittuja BI-työkaluja ovat muun muassa Microsoft Power BI, Qlik Sense ja Tableau. Kaikki edellä mainitut pystyvät tekemään monimutkaisia kyselyjä ja reaaliaikaisia analyyseja, tai niille voidaan määritellä ajastettuja toimia. Yhdenkään käyttö ei välttämättä edellytä ohjelmointiosaamista, koska ohjelmat mahdollistavat datan visuaalisen tutkimisen ja siihen kyselyjen kohdistamisen käyttöliittymän toiminnoilla.

Power BI:n yhdistettävyyys erilaisiin tietolähteisiin on jonkin verran muita parempi, kun taas Tableau sisältää valmiiksi sopivia raporttipohjia suosittuihin tietolähteisiin yhdistettäessä. Tableau:ssa on valmiita animaatioita visualisoimaan ajan kuluessa tapahtuvia muutoksia. Qlik Sense on ilmeisen kyvykäs rakentamaan kyselyitä visuaalisen käyttöliittymän kautta, kun taas Power BI:n tehokas käyttö saattaa vaatia DAX-kyselykieleen perehtymistä. (Kaur, 2021.) Microsoft Power BI on varsin yleinen ratkaisu organisaatioissa, jotka ovat muussakin toiminnassaan valinneet Microsoftin ohjelmistoekosysteemin. Samasta syystä se valittiin tämän opinnäytetyön kehitysosan toteutukseen.

3.4.1 Microsoft Power BI

Power BI:n keskeiset ominaisuudet kuten Power Query ja Power Pivot olivat alun perin Microsoft Excelin lisäosia, jotka julkaistiin Power BI-nimisenä kokonaisuutena ensimmäisenä kerran 2013 (Wikipedia 2022). Excel tunnetaan etenkin taulukkolaskentaohjelmaksi, jonka soluihin kirjoitetaan suoraan ja jonka tietolähteinä toimivat tyypillisesti yksittäiset tiedostot. Power BI:ssä tietoja käsitellään enemmän joukkona, josta voidaan laskea uusia

suureita tai suodattaa kokonaan uusia tietojoukkoja, ja joiden visualisointiin on tarjolla Exceliä paremmat työkalut.

Microsoft kuvailee Power BI -ohjelmistokokonaisuutensa koostuvan sovelluksista ja yhdistimistä (eng. connector), jotka yhdessä muuntavat eri lähteistä peräisin olevan datan ymmärrettäväksi ja interaktiiviseksi näkemykseksi (Microsoft 2022a). Lähteet voit olla toisiinsa liittymättömiä, kuten vaikka yrityksen sisäisessä käytössä oleva asiakasrekisteri ja internetissä julkaistu avoin markkinatutkimusaineisto ovat. Power BI:n ohjelmistokokonaisuus jakautuu kolmeen eri osaan. Ensimmäinen osa on Windowsin työpöytäsovellus nimeltään Power BI Desktop. Työpöytäsovellus on raporttien luomiseen tarkoitettu työkalu. Toinen osa on verkossa oleva SaaS (Software as a Service) -palvelu nimeltään Power BI -palvelu. Palveluun julkaistaan työpöytäsovelluksella luodut raportit, ja niitä käytetään kyseisessä palvelussa. Raporttien tietosisältö päivittyy suoraan palveluun, joten julkaisun jälkeen työpöytäsovellusta tarvitaan vasta, jos raportin rakennetta halutaan muuttaa. Lisäksi oma osansa on Power BI -mobiliisovellukset Windows-, iOS- ja Android-laitteille, joilla voi tarkastella ja rajatusti käsitellä raportteja. Edellä mainittujen lisäksi Microsoft tarjoaa raportointipalvelinratkaisua asiakkaille, jotka haluavat pitää kaikki tietonsa esimerkiksi oman palomuurinsa takana.

3.5 Datan visualisointi

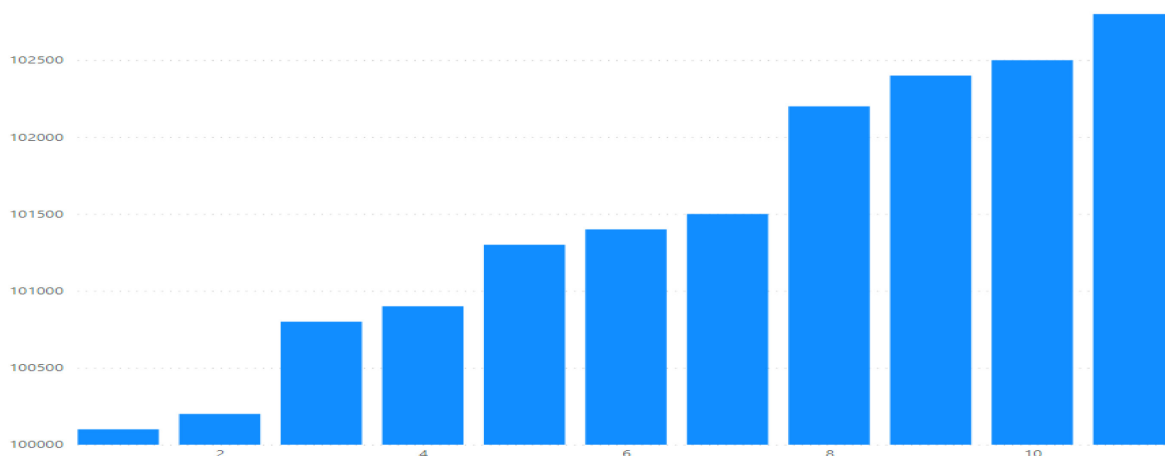
Tiedon visualisoinnilla tarkoitetaan tiedon visuaalista esittämistä sen havainnollisuuden, käsitettävyyden, selitettävyyden, löydettävyyden ja, tai, viestittävyyden vuoksi. Tieto on tyypillisesti numeerista, mutta se voi yhtä hyvin olla laadullista. Tieto itsessään on abstraktia, mutta tiedon visualisoinnilla voidaan tuoda sen erilaisia ominaisuuksia esiin. Tiedon visualisoinnilla on aina jokin tarkoitus. (Zheng 2022.) Visuaalisilla elementeillä kuten pylväillä, graafeilla ja kartoilla helpotetaan trendien, rakenteiden ja poikkeavien havaintojen löytämistä. Nykyisenä suurien datamäärien aikana visualisointi on olennainen keino analysointia ja tietoon perustuvien päätösten tekoa varten. (Tableau 2022b.)

Ihmisen kyky muodostaa kokonaiskuvaa pelkästä numeerisesta tiedosta on rajallinen. Tiedon visualisoinnin tarkoitus onkin parantaa aineiston ymmärtämistä, ja mahdollistaa sen helppo tutkiminen, selittäminen ja löydösten eteenpäin viestiminen. Visualisointi voi olla luonteeltaan raportoivaa, kuten yrityksen osavuositarkastuksessa, tai analysoivaa, kuten vaikka arvojen sijoittaminen koordinaatistoon poikkeavien havaintojen löytämiseksi. (Zheng 2022.) Se voi myös olla reaaliaikaista ja monitoroivaa. Auton mittaristoa esimerkkinä käyttäen voidaan havainnollistaa aineiston ymmärrettävyyden parannusta visualisoinnilla. Auton mittaristo koostuu useista erilaisista mittareista, kuten polttoaine- ja öljynpainemittareista. Polttoainemittari ei kerro meille kuitenkaan numeerisesti, kuinka monta litraa

polttoainetta tankissa on. Numeroita tai litramääriä ei yleensä ole merkitty mittariin ollenkaan. Sen sijaan mittaria vilkaisemalla saadaan nopeasti käsitys polttoaineen määrästä suhteessa tankin tilavuuteen sekä akuutista tankkaustarpeesta, mikäli viisari on siirtynyt punaiselle alueelle.

Öljynpainemittarin kanssa harvempi joutuu ikinä tekemisiin. Öljynpainemittari tarkkailee silti öljynpainetta jatkuvasti, mutta auton kojetaulussa sen visualisointi pysyy piilossa. Tilanne muuttuu vasta, kun sille asetetut raja-arvot ylittyvät. Tällöin öljynpaineen varoitusvalo syttyy, mutta sen tarkoitus ei ole informoida kuljettajaa siitä kuinka monta baria painetta on. Ainut tarkoitus on saada kuljettaja lisäämään öljyä. (Kujansivu ym. 2007 21.)

Tilastoja tai oikeastaan mitä tahansa numeerista tietoa visualisoidessa on mahdollista saada asiat näyttämään joltain, mitä ne eivät ole – tahattomasti tai jopa tahallisesti. Skaalaa muuttamalla saadaan havaintojen erot näyttämään tarpeen mukaan joko merkittävilta tai merkityksettömillä. Alla olevassa esimerkissä (Kuvio 4) muutos 100000:sta 102500:een saadaan näyttämään merkityksellisemmältä leikkaamalla pystyakselia. Muutos on silti vain +2,5%.

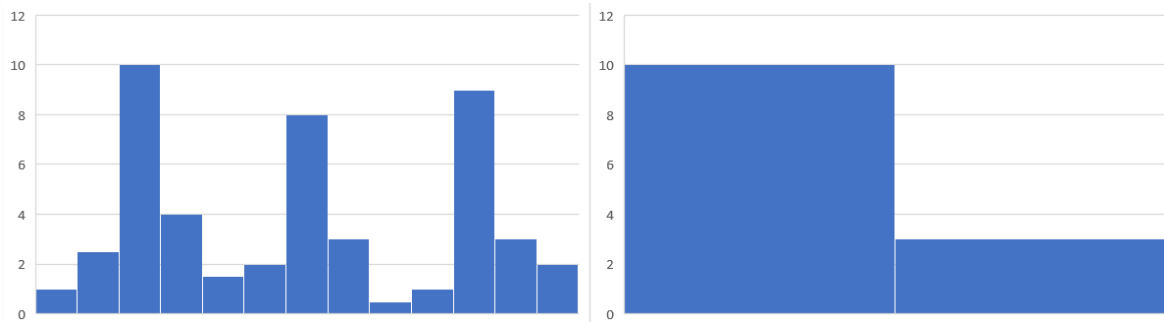


Kuvio 4. Leikattu pystyakseli

Myös kaavion koordinaatiston muotoa kaventamalla voidaan käyrän muutos saada näyttämään jyrkemmältä (Heikkinen 2019). Värien harkitsemattomalla käytöllä voidaan luoda mielikuvia suurista muutoksista, vaikka tosiasiallisesti kaksi täysin eri väriä tarkoittaisivat vain hieman eri arvoja tietolähteessä. Suhteellisesti asioita esitettäessä voidaan myös sortua virheisiin: suuressa kaupungissa voi tapahtua suhteellisesti vähemmän rikoksia kuin pienessä kylässä, mutta jos pienessä kylässä asuu vain 10 ihmistä ja siellä tehdään yksi rikos vuodessa, ei sitä ole mielekästä verrata miljoonakaupungin rikostilastoihin.

Kartoissa jonkin arvon istuttaminen maantieteelliseen alueeseen on hankalaa: onko esitetyllä pinta-alalla oikeasti jokin sidos esitettyyn arvoon (Emery 2020)? Mikäli jokainen

suomalainen kunta väritetään sen kannatetuimmalla puolueella, Suomi näyttää erittäin keskustalaiselta. Kunnan pinta-alalla ei kuitenkaan ole yhteyttä sen asukasluvun eli äänisaaliin kanssa. Dataa pyöristämällä tai liian karkeasti lokeroimalla voidaan myös piilottaa ilmiöitä, jotka näkyisivät tarkemmassa esityksessä. Alla olevassa esimerkissä (Kuvio 5) on kaksi histogrammia samasta datasta, joista jälkimmäisen tarkkuus ei riitä tuomaan ilmiötä esille.



Kuvio 5. Oikean tarkkuuden merkitys visualisoinnissa

Väärä visualisointi voi esittää koko ilmiön väärin. Esimerkiksi ympyräkaavio ilmentää aina jotain täyttä, eli sataprosenttista, mutta havainnot voivat olla yhteensä enemmän tai vähemmän kuin 100 %. Myös trendiviivojen esittäminen päällekkäin, mutta eri asteikoilla mahdollistaa toisen trendin merkittävyyden korostamisen.

Tämän opinnäytetyön kehitysosassa tulikin kiinnittää huomiota siihen, että visualisoinnit kuvasivat ilmiöitä eivät ainoastaan oikein, mutta myös totuudenmukaisesti. Kehitystyön viimeisessä versiossa ilmaistiin eri väreillä pääasiallisesti vain ajallista eroavaisuutta, mutta yhdessä visualisoinnissa värit kuvasivat myös laadullista eroa. Huonomman vaihtoehdon väri pyrittiin valitsemaan selkeäksi, mutta ei ylikorostetuksi. Viimeiseen versioon ei myöskään päätynyt muita kuin pylväsdigrammeja, joiden kaikkien akselit alkoivat nolasta.

3.6 Tietosuoja

3.6.1 Mitä on tietosuoja?

Tietosuoja on perusoikeus, joka suojaa meidän henkilötietojamme. Mikäli joku käyttää sinun henkilötietojasi, ihan mihin tahansa, hänen tulee varmistua siitä, että käyttö perustuu lakiin, ja sinun oikeutesi ja vapautesi toteutuvat tässä kyseisessä henkilötietojen käsittelyssä. Tietosuoja osoittaa miten, milloin ja millä edellytyksillä henkilötietoja voidaan käsitellä. Riippumaton viranomainen, tietosuojavaltuutettu, valvoo henkilötietojen suojaa koskevien säännösten noudattamista. (Tietosuojavaltuutetun toimisto 2021.)

Henkilötietoa on sellainen tieto, jonka perusteella henkilö voidaan tunnistaa. Tämä voi tapahtua suoraan, kuten esimerkiksi nimen tai sosiaaliturvatunnuksen avulla, mutta se voi

tapahtua myös välillisesti, kuten yhdistämällä yksittäinen tieto johonkin toiseen tietoon, joka mahdollistaa tunnistamisen. (Tietosuoja-valtuutetun toimisto 2022d.) Välillinen tunnistaminen voi tapahtua esimerkiksi tilanteessa, jossa yrityksellä on käytössään rekisteri, joka sisältää henkilön nimen yhdistettynä IP-osoitetietoihin, sekä rekisteri, joka sisältää IP-osoitetiedot yhdistettynä internet-hakujen tietoihin. Tällöin yrityksellä on käytössään rekisteri, joka sisältää henkilötiedot yhdistettynä internet-hakujen tietoihin.

Tietosuoja-asetus suojaa henkilötietoja riippumatta käytetystä tekniikasta tai säilytystavasta. Tiedot voivat olla digitaalisia, tai ne voivat yhtä hyvin olla paperimapissa tai videokuvassa. Mikäli henkilö on tiedoista tunnistettavissa, suoraan tai välillisesti, ne ovat henkilötietoja. Tämä koskee myös tietoja, jotka ovat palautettavissa tunnistettavaan muotoon, vaikka eivät sitä sillä hetkellä olisi. Tietosuojalainsäädäntöä ei sovelleta, jos käsittely on henkilökohtaista, eikä se liity ammatilliseen tai kaupalliseen toimintaan. Esimerkkinä tällaisesta käsittelystä käy muun muassa puhelimeen tallennetut ystävien puhelinnumerot ja osoitteet. (Tietosuoja-valtuutetun toimisto 2022d.)

Tietosuojan yhteydessä puhuttaessa rekisteröidystä, tarkoitetaan henkilöä, jonka henkilötiedot on tallennettu johonkin rekisteriin, eli henkilötietorekisteriin. Rekisterinpitäjä on puolestaan taho, joka henkilötietorekisteristä vastaa. REDU on rekisterinpitäjä usealle eri henkilötietorekisterille. REDUn perustehtävänä on järjestää ammatillisesta koulutuksesta ja vapaasta sivistystyöstä annettujen lakien mukaista koulutusta, kehittämistoimintaa ja oheispalveluja (REDU 2022b). Henkilötietojen käsittely on siten REDUn perustehtävän mukaista toimintaa. REDUn tietosuojapolitiikka nojaa organisaation itselleen määrittelemiin periaatteisiin:

- REDU noudattaa voimassa olevaa tietosuojalainsäädäntöä
- Se käsittelee vain sellaisia henkilötietoja ja sillä laajuudella, joihin sille on lainsäädännössä annettu oikeus tai rekisteröity on itse antanut suostumuksen
- Se kunnioittaa asiakkaidensa, kumppaneidensa ja henkilökuntansa oikeutta yksityisyyteen ja käsittelee heidän henkilötietojaan vastuullisesti ja luottamuksellisesti
- Tietosuoja varmistetaan tietoturvallisilla tietojärjestelmillä ja -palveluilla
- Koulutusorganisaationa se pyrkii toimimaan tietosuojavaikuttajana esimerkkinä ja mallina
- Se seuraa, kehittää ja parantaa toimintaansa mukaan lukien tietosuoja laadunhallintajärjestelmän jatkuvan parantamisen periaatteiden mukaisesti. (REDU 2021.)

REDUn sisäisen toimintakäsikirjan mukaisesti kaikki uudet henkilötietojen käsittelytoiminnot edellyttävät tietosuoja-asetuksen laatimista sekä tarvittaessa tietojen käsittelyn riskien

arviointia. Tämän opinnäytetyön kehitysosassa käsiteltiin henkilötietoja, ja osa käsittelytoiminnoista tunnistettiin korkeariskisiksi. REDUn tietosuojavastaavan suosituksesta tietosuoja otettiin mukaan jo uusien käsittelytoimintojen suunnitteluvaiheessa.

3.6.2 Tietosuojalainsäädäntö

Yksityisyyden suojasta säädetään Suomessa muun muassa perustuslaissa perusoikeutena, julkisuuslaissa, laissa yksityisyyden suojasta työelämässä, tietosuojalaissa sekä laissa sähköisen viestinnän palveluista. Euroopan unionin tasolla yksityisyyttä suojataan yleisellä tietosuojasetuksella 2016/697, jota sovelletaan suoraan, sekä rikosasioiden tietosuojadirektiivillä, joka on saatettu osaksi suomalaista lainsäädäntöä. (Minilex 2021.) EU:n yleinen tietosuojasetus ("GDPR") ja sitä kansallisesti tarkentava tietosuojalaki toteuttavat henkilötietojen suojaa ja näin osaltaan myös yksityisyyden suojaa. EU:n tietosuojasetus on ollut voimassa kaikissa EU:n jäsenvaltioissa 25.5.2018 alkaen.

Kansallinen tietosuojalaki 1050/2018 ei muodosta itsenäistä ja kattavaa sääntelykokonaisuutta, vaan sitä sovelletaan rinnakkain edellä mainitun EU:n yleisen tietosuojasetuksen kanssa. Tietosuojalaissa säädetään muun muassa kansallisesta tietosuojaviranomaisesta (tietosuojavaltuutettu) sekä eräistä henkilötietojen käsittelyyn liittyvistä erityistilanteista kuten sananvapaudesta ja henkilötietojen suojan yhteensovittamisesta. Tietosuojasetus ja tietosuojalaki koskevat lähtökohtaisesti kaikenlaista henkilötietojen käsittelyä. Ne sisältävät säännökset rekisteröidyn oikeuksista sekä rekisterinpitäjän ja henkilötietojen käsittelijän velvollisuuksista. (Finlex 2018.)

EU:n yleinen tietosuojasetus on maailmanlaajuisesti tällä hetkellä tiukin tietosuojaa määrittelevä laki tai asetusta. Vastaavia mutta miedompia säädöksiä ovat muun muassa Kalifornian osavaltiossa "California Consumer Privacy Act" (CCPA), Brasilian "Lei Geral de Proteção de Dados" (LGPD) sekä Kanadan "Digital Charter Implementation Act". Henkilötietosuojan merkitystä aletaan nykyisin näkemään ei pelkästään yksilöiden, vaan myös organisaatioiden etuna (Titus, 2021).

3.6.3 Rekisteröidyn oikeudet

Tietosuojan turvaamia rekisteröidyn oikeuksia on kahdeksan erilaista. Rekisteröidyllä on *oikeus saada tietoa henkilötietojensa käsittelystä*. Hänelle tulee toisin sanoen kertoa, miten ja miksi hänen henkilötietojaan kerätään, ja mihin niitä käytetään. Käsittelyn tulee olla läpinäkyvää. Informaatio on tarjottava rekisteröidylle tiiviissä ja helposti ymmärrettävässä muodossa. Lisäksi *rekisteröidyllä on oikeus saada pääsy tietoihinsa*. Jotta rekisteröity voi arvioida ja varmistaa käsittelyn lainmukaisuuden, hänellä on oikeus saada rekisterinpitäjältä

vahvistus siitä, käsittelee tämä hänen henkilötietojaan. Mikäli käsittelee, rekisteröidylle on pyynnöstä toimitettava jäljennös käsiteltävistä henkilötiedoista, sekä muun muassa tieto käsittelyn tarkoituksesta, siitä mistä tiedot on saatu, sekä niiden suunniteltu säilytysaika. Lisäksi *rekisteröidyllä on oikeus oikaista tietoja*. Mikäli henkilötiedot ovat epätarkkoja, virheellisiä tai puutteellisia, rekisteröidyllä on oikeus vaatia niiden oikaisua. Lisäksi *rekisteröidyllä on tietyissä tilanteissa oikeus tulla unohdetuksi*. Tällainen tilanne on esimerkiksi jos rekisteröity peruuttaa suostumuksensa, johon käsittely on perustunut. Mikäli käsittelylle ei ole enää muuta perustetta, käsittelyn tulee loppua ja henkilötieto tulee poistaa. Lisäksi *rekisteröidyllä on oikeus rajoittaa tietojen käsittelyä*. Tällainen tilanne on esimerkiksi, jos rekisteröidyllä ja rekisterinpitäjällä ovat eri mieltä henkilötietojen paikkansapitävyydestä. Tällöin käsittelyä rajoitetaan ajaksi, jonka kuluessa tietojen oikeellisuus varmistetaan. Lisäksi *rekisteröidyllä on oikeus siirtää tiedot järjestelmästä toiseen*, mikäli se on teknisesti mahdollista. Tätä sovelletaan ainoastaan henkilötietojen automaattiseen käsittelyyn. Lisäksi rekisteröidyllä on oikeus saada hänen itse toimittamansa henkilötiedot takaisin jäsennellyssä, yleisesti käytetyssä ja koneellisesti luettavassa muodossa. Lisäksi *henkilöllä on tietyissä tilanteissa oikeus vastustaa henkilötietojensa käsittelyä* eli pyytää, että niitä ei käsiteltäisi lainkaan. (Tietosuojavaltuutetun toimisto 2022e.) Tällainen tilanne on esimerkiksi, jos rekisteröity vastustaa Verohallintoa luovuttamasta hänen tuloverotietojansa tiedotusvälineille journalistiseen tarkoitukseen (Verohallinto 2021). *Henkilöllä on lisäksi oikeus olla joutumatta automaattisen päätöksenteon kohteeksi*. Mikäli päätöksenteolla on oikeusvaikutuksia, tai päätöksenteko vaikuttaa rekisteröityyn merkittävästi, hänellä on oikeus vaatia, että päätöksen tekee ihminen. Kieltoon on poikkeuksia. (Tietosuojavaltuutetun toimisto 2022a.)

3.6.4 Tietosuoja ja tietoturva

Tietosuoja keskittyy siihen, miten henkilötietoa kerätään ja käytetään, ja miksi näin toimitaan. Nykyaikaisessa, pitkälle digitalisoituneessa maailmassa rekisteröidyn oikeuksien toteutuminen vaatii kuitenkin muutakin kuin lakeja ja asetuksia. Tietoturvaa ovat ne teknologiset ja hallinnolliset ratkaisut, jotka turvaavat dataa ulkoisilta ja sisäisiltä uhkilta. Organisaation perusteelliset tietoturvakäytännöt kattavat niin ihmiset, prosessit kuin teknisetkin ominaisuudet. (Titus 2021.)

Pelkän tietoturvan toteuttaminen ei siis toteuta tietosuojaa. Tietoturva suojelee organisaatiota vaikkapa kyberhyökkäykseltä, jolla pyritään varastamaan organisaation hallussa olevia henkilötietoja. Tietosuoja taas huolehtii siitä, että organisaatio suhtautuu näihin hallussaan oleviin henkilötietoihin asian vaatimalla vakavuudella, ja käsittelee niitä ainoastaan lain määrittämissä rajoissa. Toisin sanoen, kun yritys toteuttaa tietosuojaa, se huolehtii siitä, että arat tiedot, kuten henkilötiedot, ovat muun muassa tietoturvan turvaamia.

3.6.5 Henkilötietojen käsittelyn perusteet ja oikeutetun edun arviointi

Henkilötietojen käsittely edellyttää aina lakiin kirjattua perustetta. Peruste on määriteltävä, ennen kuin käsittely voidaan aloittaa. Kun jollekin käsittelytoiminnolle on määritelty peruste, ei sitä voi enää vaihtaa toiseen aloittamatta koko prosessia alusta. Käsittelyperuste vaikuttaa siihen, mitkä rekisteröidyn oikeudet (kappale 3.6.3) voidaan toteuttaa. Tietosuoja-asetuksessa on kuusi erilaista perustetta, joilla henkilötietojen käsittely voidaan perustella. Ensimmäinen niistä on *suostumus*. Rekisteröity voi antaa suostumuksensa henkilötietojensa käsittelyyn yhtä tai useampaa tarkoitusta varten. Suostumuksen on oltava vapaaehtoinen, yksilöity tahdonilmaisu, joka rekisterinpitäjän on pystyttävä todistamaan olemassa olevaksi. (Tietosuojavaltuutetun toimisto 2022e, Tietosuojavaltuutetun toimisto 2022i.) Esimerkiksi selaimen ”Hyväksyn evästeet” -ponnahdusikkuna on tässä tarkoitettu suostumus henkilötietojen käsittelyyn (Innowise 2019). Toinen mahdollinen käsittelyperuste on *sopimus*. Kun rekisteröity on osapuolena sopimuksessa, hänen henkilötietojaan voidaan käsitellä sopimuksen täytäntöönpanemiseksi. Jos esimerkiksi asiakas tilaa verkkokaupasta tuotteita, hänen henkilötietojaan täytyy käsitellä jotta tuotteet voidaan toimittaa hänelle. Lisäksi käsittelyperusteena voi olla *lakisääteinen velvoite*. Tämä voi koskea niin yksityisellä kuin julkisella sektorilla toimivaa yritystä tai organisaatiota, jolla on Euroopan unionin tai sen jäsenvaltion lakiin perustuva velvoite käsitellä henkilötietoja. Esimerkiksi työnantajan ilmoittaessa palkkatiedot veroviranomaiselle on kyse lakisääteisestä velvoitteesta. Myös *elintärkeiden etujen suojaaminen* voi käydä käsittelyperusteeksi. Kriisitilanteissa kuten luonnonkatastrofeissa tai epidemioissa voi henkilötietojen käsittely olla perusteltua esimerkiksi epidemian leviämistä seurattaessa. Käsittelyn perusteena voi toimia myös *yleinen etu ja julkinen valta*. Henkilötietoja saa käsitellä, mikäli rekisterinpitäjälle myönnetyn julkisen vallan käyttäminen sitä edellyttää. Julkinen valta täytyy olla lakiin tai säädökseen kirjattu. Tieteellinen tutkimus tai tilastointi käyvät esimerkkeinä yleistä etua palvelevista toiminnoista, joissa henkilötietoa saa käsitellä. Lisäksi on niin kutsuttu *rekisterinpitäjän oikeutettu etu*. Henkilötietoja voidaan käsitellä, mikäli se on tarpeen rekisterinpitäjän niin kutsuttujen oikeutettujen etujen toteuttamiseksi. Tätä etua tulee kuitenkin punnita rekisteröidyn perusoikeuksia vastaan. Tätä punnitusta kutsutaan tasapainotestiksi. Oikeutettu etu voi olla olemassa esimerkiksi silloin, kun rekisteröity on työsuhteessa rekisterinpitäjään. (Tietosuojavaltuutetun toimisto 2022e.)

Tämän opinnäytetyön kehitysosassa käytettyjen henkilötietojen käsittely perustui rekisterinpitäjän oikeutettuun etuun. Kehitystyön tavoitteena oli kehittää tiettyjen henkilöstöresurssin tietojen raportointia. Oikeutettu etu tässä tapauksessa tarkoittaa sitä, että koska rekisteröity on rekisterinpitäjään työsuhteessa, on rekisterinpitäjällä ”oikeus” tulla raportoiduksi työntekijänsä poissaoloista ja siitä, miten rekisteröity käyttää hänelle annettua resurssia eli osaamisen kehittämisen resurssia.

3.6.6 Vaikutustenarviointi ja riskianalyysi

Kun suunnitellaan henkilötietojen käsittelyn aloittamista, oli kyse sitten uusista tiedoista, tai pelkästään uudesta tavasta käsitellä vanhoja tietoja, tulee rekisterinpitäjän arvioida käsitteilyyn liittyviä riskejä. Tavoitteena on tunnistaa uhat ja reagoida niihin jo suunnitteluvaiheessa, jotta rekisteröityjen oikeuksiin kohdistuvat riskit ovat mahdollisimman hyvin tiedossa ja niitä voidaan pyrkiä hallitsemaan ja minimoimaan. Riskiarvio tehdäänkin nimenomaan rekisteröidyn näkökulmasta, ja sen tulee pureutua myös potentiaalsiin haittoihin, joita rekisteröidylle koituisi riskien toteutumisesta. Henkilötietojen vuotaminen voi tarkoittaa rekisteröidylle esimerkiksi sitä, että hänen nimissään saatetaan tehdä petos, mikä aiheuttaisi huomattavaa taloudellista haittaa. Se voi tarkoittaa myös mainehaittaa, jolla saattaa olla henkilön loppuelämään vaikuttavia seurauksia. Riskiarviota tehdessä on pyrittävä saamaan mahdollisimman kattava ja selkeä käsitys siitä, mitä, miten ja miksi henkilötietojen käsittelyä ollaan aloittamassa. Siinä tulee huomioida käsittelyn luonne, laajuus, asiayhteys ja tarkoitukset. Kun kaikki potentiaaliset riskit on tunnistettu, tulee arvioida riskin ja siitä aiheutuvan haitan vakavuutta ja toteutumisen todennäköisyyttä. Riskien arviointia ei tehdä pelkästään käsittelytoimintoja kehitettäessä. Toimien riittävyttä suhteessa riskiin tulee arvioida jatkuvasti ja sitä on päivitettävä tarvittaessa. Rekisterinpitäjällä on osoitusvelvollisuus riskiperusteisen lähestymistavan noudattamisesta. (Tietosuojavaltuutetun toimisto 2022f, Tietosuojavaltuutetun toimisto 2022k.)

3.6.7 Tasapainotesti

Koska opinnäytetyössä kehitettävän uuden henkilötiedon käsittelytoiminnon käsittelyperusteen katsottiin olevan ”rekisterinpitäjän oikeutettu etu”, tuli tietosuoja-asetuksen mukaisesti käsittelytoiminnosta laatia niin kutsuttu tasapainotesti. Tasapainotesti arvioi, käykö oikeutettu etu käsittelyn perusteeksi. Rekisterinpitäjä on velvoitettu käyttämään tasapainotestiä punnitakseen omia etujaan rekisteröidyn etuja ja oikeuksia vastaan, ja siitä tulee laatia kirjallinen kuvaus, jolla rekisterinpitäjä voi osoittaa toiminnan olevan tietosuoja-asetuksen mukaista.

3.6.8 Henkilötieto ja anonymisointi

Henkilötietoja ovat kaikki tunnistettuun tai tunnistetavissa olevaan luonnolliseen henkilöön liittyvät tiedot. Myös tiedot, jotka ovat erillään, mutta yhdistettynä mahdollistavat tunnistamisen, ovat henkilötietoja. Salatut, anonymisoidut tai pseudonymisoidut henkilötiedot ovat edelleen henkilötietoja, mikäli anonymisointi tai vastaava prosessi voidaan kääntää. Anonymisoinnin on siis oltava peruuttamatonta. (Euroopan komissio 2022.) Jos esimerkiksi

sairaala merkitsee potilaan vuoteeseen symbolin, ja kirjanpidossa on vastaava symboli, on henkilörekisteri syntynyt. Symbolien käyttö nimien sijaan ei muuta asiaa. Tai, jos tutkimuskäyttöön hankitun henkilötietorekisterin nimiä yritetään anonymisoida, mutta anonymisointi tehdään funktiolla, joka käyttää syötteenään henkilön nimeä. Funktio saattaa olla käännettävissä, jolloin anonymisoinnin peruuttamattomuudesta ei voida olla enää varmoja. Kyseessä ovat siis edelleen henkilötiedot.

3.6.9 Erityiset henkilötietoryhmät ja arkaluonteiset tiedot

Niin sanottuihin erityisiin henkilötietoryhmiin kuuluvien henkilötietojen käsittely on lähtökohteisesti aina kiellettyä. Näistä tiedoista saa selville joko henkilön rodun, etnisen alkuperän, poliittisen suuntautuneisuuden, uskonnollisen tai filosofisen vakaumuksen, ammatiliiton jäsenyyden, terveyttä koskevia tietoja, seksuaalisen suuntautumisen tai käyttäytymisen, tai geneettisiä tai biometrisiä tietoja henkilön tunnistamiseen. Näiden tietojen yksityisyyttä tulee varjella tarkasti, sillä niiden käsittely aiheuttaa aina vähintään jonkinlaisia riskejä, jotka kohdistuvat henkilön perusoikeuksiin ja -vapauksiin. (Tietosuojavaltuutetun toimisto 2022b.)

Tämän opinnäytetyön kehitysosa koski muun muassa sairaspotilaiden raportoinnin kehittämistä. Pelkkä tieto sairaslomalla olemisesta ei kuitenkaan yksinään ole henkilötietolain tarkoittama arkaluonteinen henkilötieto (Tietosuojavaltuutetun toimisto 2020, 50). Mikäli tieto sisältäisi myös syyn sairaspotilaaloon, muuttuisi tieto arkaluonteiseksi.

3.6.10 Henkilötietojen käsittely

Henkilötietojen käsittelyllä tarkoitetaan mitä tahansa toimintaa, jonka osallisena henkilötiedot ovat. Se voi siis olla tietojen keräämistä, mutta yhtä lailla säilyttämistä, käyttämistä, siirtämistä tai luovuttamistakin. Kaikki toiminta, joka kattaa ajan aina käsittelyn suunnittelusta henkilötietojen lopulliseen poistamiseen on henkilötietojen käsittelyä. (Tietosuojavaltuutetun toimisto 2021.)

3.6.11 Henkilörekisteri

Henkilörekisteri on jäsennelty tietojoukko, joka sisältää henkilötietoa. Tietojoukko on rekisteri, mikäli sen tiedot ovat tavalla tai toisella saatavissa, ja riippumatta siitä, onko se keskitetty, hajautettu, tai vaikka jaettu maantieteellisin perustein. Myöskään tallennusmedialla ei ole merkitystä. Myös yksittäisen henkilön henkilötietoja on käsiteltävä tietosuojasäädösten mukaisesti. Tämä koskee myös henkilörekistereistä otettuja henkilön tai henkilöiden tietoja sisältävää otetta, kuten esimerkiksi ladattua tiedostoa tai tulostetta. (Minilex 2022.)

3.6.12 Suppeat ja laajat pysyvät henkilötietorekisterit

REDUssa henkilöresterit ja niitä koskevat ohjeistus on jaettu kolmeen luokkaan. Ensimmäinen luokka on ”laajat ja pysyvät henkilöresterit”, toinen on ”suppeat ja tilapäiset henkilöresterit” ja kolmas on ”tilapäiset apuresterit”. Laajalla ja pysyvällä henkilöresterillä tarkoitetaan rekisteriä, joka sisältää lukumääräisesti yli 50:n henkilön tietoja, ja jota on tarkoitus ylläpitää toistaiseksi. Tämän opinnäytetyön kehitystyön ideoima käsittelytoiminto rekistereineen kuuluu tähän luokkaan.

Suppeat ja tilapäiset henkilöresterit tarkoittavat REDUssa alle 50:n henkilön tietoja sisältäviä rekistereitä, joiden käyttötarkoitus on tilapäinen, ja joka on hyvin pienen ja rajatun joukon käytössä, esimerkiksi opettajan tai tietyn tiimin opettajien. Apuresteriksi REDUssa kutsutaan laajasta rekisteristä kopioituja otteita. Tällaisia ovat esimerkiksi suodatetut tiedot opiskelijahallintojärjestelmästä, jotka on ladattu tiedostoon tai tulostettu.

3.6.13 Käsittelytoimintojen luettelo

Mikäli organisaatiossa on yli 250 työntekijää, tulee organisaation ylläpitää selostetta kaikista henkilötietojen käsittelytoiminnoistaan. Seloste tulee myös tehdä työntekijöiden määrästä riippumatta, jos käsiteltävät tiedot ovat arkaluonteisia. (Tietosuojavaltuutetun toimisto, 2022m.) Isona organisaationa REDU on velvoitettu käsittelytoimintojen dokumentointiin. Rekisteröityjen informointivelvollisuus sekä käsittelytoimintojen selosteen laadintavelvoite hoidetaan käsittelytoiminnoittain laadituilla ja laadittavilla tietosuojaselosteilla, jotka tallennetaan keskitetysti organisaation sisäiseen verkkoon. Lisäksi niiden käsittelytoimintojen selosteet joissa käsitellään ulkoisten asiakkaiden tietoja, ovat jakelussa myös julkisen verkon puolella REDUn sivuilla. Käsittelytoimintojen luettelo on organisaation sisäinen asiakirja, jonka on tarkoitus toimia apuvälineenä henkilötietojen käsittelyn hahmottamiseen. Omalta osaltaan sen on tarkoitus osoittaa, että organisaatio käsittelee henkilötietoja tietosuojalainsäädännön mukaisesti.

3.6.14 Tietosuojaselosteet

Edellä mainittu käsittelytoimintojen luettelo on rakennettu tietosuojaselosteista. Tietosuojaseloste on tiivistelmä henkilötietojen käsittelystä yhden rekisterin osalta. Tietosuojaselostetta ei ole välttämättä tarkoitettu suoraan rekisteröidyn informointiin, joskin sitä voidaan käyttää siihen, mutta ennen kaikkea sitä voidaan hyödyntää rekisteröidylle suunnatun informaation tuottamisessa. Tietosuojaselosteeseen kirjataan rekisteristä vastaavan organisaation tiedot, henkilötietojen käsittelyn tarkoitus ja oikeutus, sen tietosisältö, tiedon lähde, tiedon vastaanottaja, ja muun muassa tietojen säilytysaika. Lisäksi siihen kirjataan

rekisteröityjen oikeudet. Tietosuojaselosteiden tulisi olla ymmärrettäviä ja helposti saavutettavia. (Agendum 2019.)

3.6.15 Osoitusvelvollisuus

Tietosuoja-asetuksen keskeinen periaate on ns. osoitusvelvollisuus. Osoitusvelvollisuus tarkoittaa, että rekisterinpitäjä pystyy osoittamaan tietosuojalainsäädännön noudattamisen. Sen tarkoituksena on näyttää, että rekisterinpitäjä kunnioittaa henkilötietojen käsittelyn kohteena olevien tietosuojaa, ja että se aktiivisesti toimii tunnistaakseen tietosuojaan liittyviä riskejä. (Tietosuojavaltuutetun toimisto 2022f.) Osoitusvelvollisuutta toteutetaan eri tavoin. Esimerkiksi aiemmin mainitut käsittelytoimintojen luettelot ja tietosuojaselosteet ovat yksi näkyvä osa osoitusvelvollisuutta, joilla rekisterinpitäjä kertoo rekisteröidyille, mitä, miten ja miksi heidän tietojaan käsitellään. Tai, jos organisaatio huomaa henkilötietojen vuodon esimerkiksi tietoturvaloukkauksen yhteydessä, voi rekisterinpitäjä osoitusvelvollisuuden avulla näyttää, mitä toimia oli tehty tällaisen tietoturvaloukkauksen riskin pienentämiseksi ja min-kälaisilla toimenpiteillä henkilötietojen suoja oli tehty. Mikäli rekisterinpitäjä ei pysty osoittamaan että toimenpiteet olisivat olleet tietosuoja-asetuksen mukaisia, voi siitä olla seurauksena mainehaitan lisäksi myös hallinnollisia seuraamuksia.

4 Kehittämisprosessin toteutus

4.1 Tutkimuksellinen kehittämistyö prosessina

Tutkimuksellisen kehittämistyön prosessi voidaan jakaa kronologisesti kuuteen eri vaiheeseen tai osa-alueeseen. Ensimmäinen vaihe oli kehittämiskohteen tunnistaminen ja alustavien tavoitteiden määrittäminen, mikä toteutui opinnäytetyösuunnitelmassa ja sitä seuranneissa keskusteluissa kehitystyöryhmän kesken. Tämän jälkeen opinnäytetyöntekijä perehtyi kehittämiskohteeseen teoriassa ja käytännössä. Kehitystyössä käytettäviin työkaluihin eli Power BI-ohjelmistoon ja M- ja DAX- kaava- sekä kyselykieliin tutustumisen jälkeen opinnäytetyöntekijä pyrki kasvattamaan ymmärrystään niin henkilöstöhallinnon kuin potilastietojärjestelmienkin uusimmista tuulista muun muassa webinaareihin osallistumalla ja lähdekirjallisuutta tutkimalla. Tietosuojaan riskianalyysin ja vaikutustenarvioinnin pohjatyönä opinnäytetyöntekijä opiskeli myös tietosuojaa syvällisemmin.

Kolmannessa vaiheessa kehittämistehtävä määriteltiin ja kehittämiskohde rajattiin. Rajauksen myötä kehitystyöhön sisällytettäväksi raporteiksi jäivät sairauspoissaolojen ja osaamisen kehittämisen resurssin raportit. Kehitystyön ulkopuolelle rajattuja raportteja olivat muun muassa henkilöstön määrää ja laatua esittelevä raportti, opettajien pätevyyttä esittelevä raportti, palvelussuhteiden päättymisen syitä ja määriä esittelevä raportti sekä henkilöstön palkkakustannuksia per opiskelija tai opintopiste esittelevä raportti.

Kehitystyöryhmän kanssa käydyissä keskusteluissa ideoitiin, minkälaista Power BI -raportointia ja visualisointia ensimmäisessä versiossa kannattaisi pyrkiä tuottamaan, sekä miten niiden jatkojalostus toteutettaisiin. Erilaisia raportteja sovittiin aluksi toteutettavaksi mahdollisimman monipuolisesti, joita voitaisiin sitten karsia ja jäljellejääneitä kehittää. Kehitystyön eri iteraatioita sovittiin tarkasteltaviksi kehitystyöryhmän kesken aina, kun uusi ja päivitetty versio olisi valmiina kommentoitavaksi. Samaan aikaan alkoi myös ilmetä tietosuojaan merkitys kokonaistyön toteuttamisen kannalta, koska kehitystyön alku- ja loppuvaiheessa käsiteltäisiin henkilötietoja, joista osa saattaisi olla korkeariskisiä käsittelytoimintoja. Tietosuojaan suunnittelu otettiin siten kehittämistehtävään mukaan.

Neljännessä vaiheessa laadittiin tietoperusta, joskin hyödyllisiä lähteitä löytyi myöhemmin lisää myös kehitystyön aikana. Tietosuojaan ja datan käsittelyyn liittyvät aiheet koettiin tärkeiksi etenkin työn käytännön toteutuksen kannalta. Henkilöstöhallinnon ja -resursoinnin sekä aineettoman pääoman lähdekirjallisuus tarjosivat laajempaa näkemystä työn kontekstiin. Tiedolla johtaminen ei ole käsitteenä kovin selkeä, mutta sen määrittelyn yrittäminen tarjosi kokonaisvaltaista näkemystä siihen, mitä tietoperustan tulisi sisältää. Käytettävien ohjelmistojen opetteluun opinnäytetyöntekijä suosi ilmaisia verkkokursseja kirjojen sijaan.

Lähestymistavaksi oli valikoitunut konstruktiiivinen tutkimus jo varhaisessa vaiheessa. Opinnäytetyösuunnitelmassa tavoitteeksi määriteltiin toimivien Power BI-raporttien tuottaminen, jotka olisivat siis opinnäytetyön konstruktioita. Viides vaihe oli kehittämishankkeen toteuttaminen, joka käsitellään seuraavissa kappaleissa 4.2 ja 4.3. Kuudes vaihe eli kehittämisprosessin ja lopputulosten arviointi ovat kappaleessa 5.

Koska kehitystyössä käsiteltiin REDUn työntekijöiden henkilötietoja, haettiin opinnäytetyölle REDUn tutkimuslupa. Tutkimusluvassa määriteltiin muun muassa salassa pidettävät tiedot sekä käsiteltävät henkilötiedot. Salassa pidettäväksi tiedoksi määriteltiin kehitettävien työkalujen tai raporttien tuottama informaatio REDUn työntekijöistä, esihenkilöistä tai toimialoista. Sen sijaan raporttien ja visualisointien rakenne, toiminta ja informaatio siitä, miten raportit toimivat, määriteltiin tiedoiksi, joita ei pidetä salassa. Käsiteltäviksi henkilötiedoiksi määriteltiin työntekijän ja toimialan osaamisen kehittämisen toteutunut tai toteutumaton resurssi, joista toimiala ja työntekijä olivat yksilöitävissä, sekä työntekijän ja toimialan terveysperustaisen poissaolon tiedot, joista toimiala ja työntekijä olivat yksilöitävissä.

4.2 Tietosuojaan toteutus

REDUssa on sisäisessä käytössä työkalu, jolla kuka tahansa työntekijä voi arvioida uuden käsittelytoiminnon tietosuovaikutuksia vaikka jo hankinnan suunnitteluvaiheessa. ”Tietosuojaan alkukartoitus ja tietosuovariskien pika-itsearviointi” - nimisellä testillä uutta toimintoa suunnitteleva henkilö pystyy tunnistamaan, että käsitelläänkö henkilötietoja ylipäättään, ja sisältyykö niiden käsittelyyn riskejä. Mikäli pika-arvioinnin tulos on se, että käsittelyyn sisältyy korkeita riskejä, tulee uudelle toiminnolle tehdä perusteellinen tietosuovaikutusten arviointi ja siihen sisältyvä riskien arviointi ennen uuden käsittelytoiminnon käytön aloittamista. Toinen vaihtoehto on suunnitella aiottu käsittelytoiminto uudelleen siten, että riskejä poistetaan tai vähennetään niin paljon, etteivät korkean riskin kriteerit enää täyty. Mikäli uuden käsittelyn toiminta ei juurikaan tästä riskien poistamisesta kärsi, se olisi joka tapauksessa pitänyt suunnitella pienimmän mahdollisen haitan kautta.

Tämän opinnäytetyön kehitysosassa kehitetty työkalu, joka käsitteli henkilöstön osaamisen kehittämiseen käyttämää resurssia, tunnistettiin henkilötiedoksi, mutta käsittelytoimintoon ei pika-arvion mukaan sisältynyt korkeita tietosuovariskejä. Tietosuojalainsäädännön toteutumiseksi sen käyttöönotto edellyttäisi tietosuojaselosteen laatimista sekä rekisterin vastuuhenkilöiden määrittelyä. Opinnäytetyössä kehitetty toinen työkalu, joka käsitteli henkilöstön sairauspoissaolotietoja, tunnistettiin pika-arviolla henkilötiedoksi ja mahdollisesti korkeita riskejä sisältäväksi käsittelytoiminnoksi. Mahdollisia korkeita riskejä aiheuttavalle käsittelytoiminnolle tehtiin tietosuojavaltuutetun toimiston ohjeistuksen mukaisesti laajempi riskianalyysi ja tietosuovaikutusten arviointi (TVA) (Tietosuojavaltuutetun toimisto 2022k). Niillä

arvioitiin käsittelyn tarpeellisuutta, oikeasuhteisuutta ja henkilötietojen käsittelystä aiheutuvia riskejä sekä tarvittavia toimenpiteitä, joilla riskeihin puututtaisiin. Tavoitteena oli arvioida ja arvioinnin tuloksen perusteella päättää, oliko jäljelle jäänyt riski oikeutettu ja hyväksyttävissä. Mikäli vaikutustenarvioinnin tulos olisi osoittanut, että riskeihin puuttumisen jälkeenkin riskitaso olisi ollut edelleen korkea, olisi asiassa tullut kuulla tietosuojavaltuutettua (viranomaista) tietosuojasetuksen 36 artiklan mukaisesti (EU:n yleinen tietosuojasetus 2016/679, 36 §).

4.2.1 Riskianalyysin ja vaikutustenarvioinnin kehitys työn aikana

REDUn sisäisen ohjeistuksen mukaisesti työlle haettiin lausunto REDUn tietosuojavastaavalta jo opinnäytetyösuunnitelman valmistuttua. Tietosuojavastaava suositteli tietosuojan riskien ja vaikutusten arvioinnin mukaan ottamista jo työn määrittelyvaiheessa, ja sen mukana pitämistä kehittämisprosessin eri vaiheissa. Kuvioon 6 on kuvattu prosessi tietosuojan kannalta.



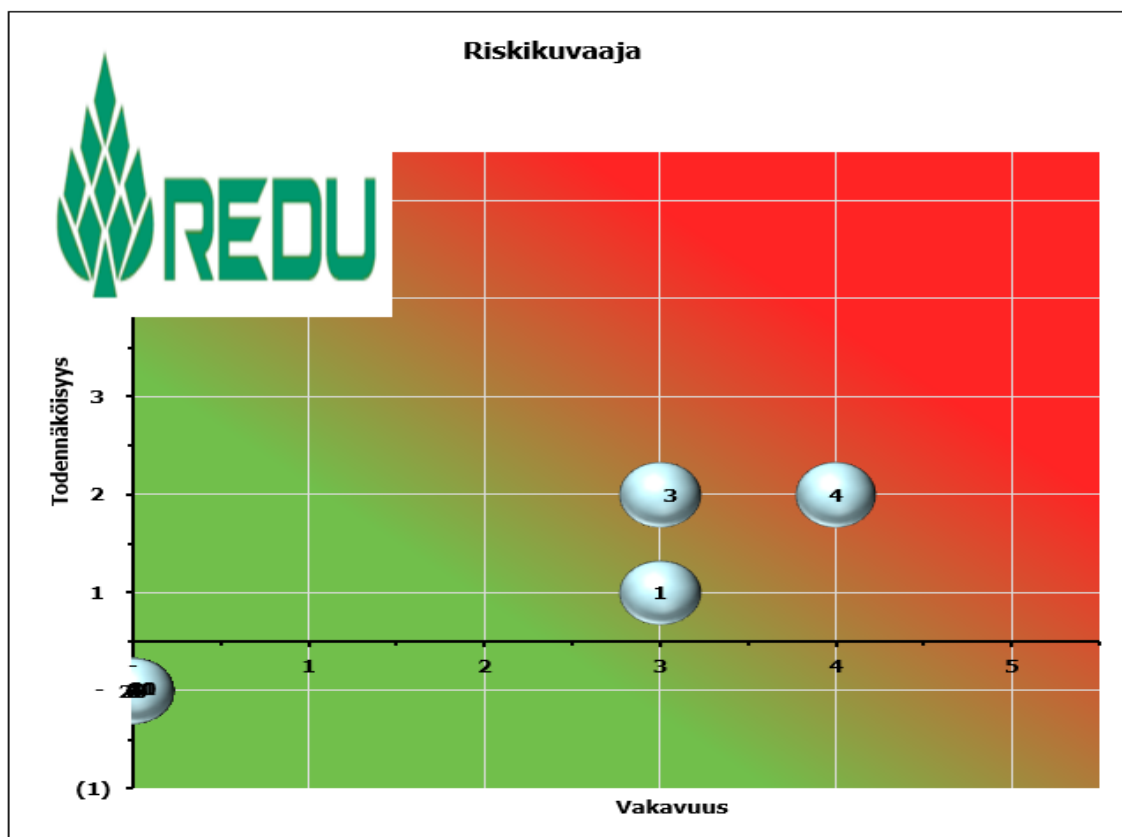
Kuvio 6. Kehitystyön prosessi tietosuojan kannalta

Tietosuojaohjeistuksen mukaisesti tehdyssä alkukartoituksessa toinen uusista käsittelytoiminnoista tunnistettiin vähäisen riskin toiminnoksi, mutta toinen tunnistettiin mahdollisesti korkean riskin aiheuttavaksi toiminnoksi. Tämä potentiaalisen korkean riskin aiheuttava toiminto käsitteli tai käsittelisi henkilöstön sairauspoissaolotietoja. Tietosuojasetuksen 35 artiklan 1 kohta edellyttää, että vaikutustenarviointi tehdään silloin, kun käsittelystä todennäköisesti seuraa korkea riski rekisteröidyn oikeuksille ja vapauksille.

Riskin todennäköisyyden arvioimiseksi on olemassa apuvälineitä, kuten REDUn sisäisessä käytössä oleva riskianalyysityökalu. Työkalun avulla on mahdollista tunnistaa uuden käsittelytoiminnon riskien vaikuttavuudet ja todennäköisyydet, ja näin ollen muodostaa käsitys kokonaisriskistä. Työkalu toimii periaatteella, jossa aluksi käydään kohta kohdalta läpi tietosuojalainsäädännön määrittelemät vaatimukset, ja se miten vaatimus täyttyy uuden käsittelytoiminnon osalta. Isoja osa-alueita on kolme: suunnitellun käsittelytoiminnon

järjestelmällinen kuvaus, henkilötietojen käsittelyn tarpeellisuuden ja oikeellisuuden arviointi sekä rekisteröidyn oikeuksille aiheutuvien riskien arviointi. Vaatimukset on eritelty hyvin seikkaperäisesti, kuten esimerkiksi käsittelytoiminnon arkkitehtuurin, tietovuon ja teknisten sekä organisatoristen toimien kuvaukset. Vaatimus voi olla esimerkiksi "Tietojen säilytysaika on määritelty niiden käyttötarkoituksen mukaisesti ja niiden poistaminen säilytysajan päätyttyä toteutetaan (myös varmuuskopiot) ja tiedon koko elinkaaren hallinta suunnitellaan jo alussa". Mikäli vaatimus ei täyty, siitä on tunnistettu riski joka kirjataan.

Edellä mainitulla tavalla kootut riskit viedään edelleen riskianalyysityökaluun. Riskin kuvauksen lisäksi määritellään seuraukset riskin toteutuessa, kuten esimerkiksi "Henkilöön yhdistettävän sairauspoissaolotiedon vuotaminen, joka koskee satoja ihmisiä". Vaikuttavuus määritellään ennalta annettujen reunaehtojen avulla, joihin vaikuttavat esimerkiksi käsiteltävät henkilötietoryhmät, rekisteröidyn mahdollisuus käyttää oikeuksiaan riskin toteutumisen jälkeen, potentiaaliset taloudelliset vahingot tai kuinka pitkistä vahingoista ajallisesti puhutaan. Riskin todennäköisyys määritellään asteikolla mitätön, epätodennäköinen, todennäköinen, erittäin todennäköinen ja odotettavissa oleva. Vaikuttavuus- ja todennäköisyys kerrotaan keskenään, jolloin saadaan jokaiselle riskille oma riskilukunsa. Riskiluvut siirtyvät edelleen riskikuvaajalle (Kuva 1).



Kuva 1. REDUn riskianalyysityökalun riskikuvaaja.

Tunnistetuille riskeille määritellään samaan dokumenttiin myös hallintatoimenpiteet, jotka pyrkivät pienentämään riskiä. Määritetyille hallintatoimenpiteille olisi suotavaa myös määritellä lisäresurssin tarve, mikäli sellainen on, ja aikataulu, vastuuhenkilöt, valmiusaste sekä riskin hallintatoimenpiteen seurannan ja raportoinnin toteutus. Tunnistettuja riskejä olivat muun muassa pilvipalvelujen tietoturva, joissa ulkoistetun palvelun tietosuojan tasoa voi olla hankala arvioida. Vaikka palvelu itsessään olisi hyvinkin turvallinen, voivat rekisterinpitäjän päässä tapahtuvat konfiguraatiovirheet tai puutteet tietotaidossa aiheuttaa vuotoja. Tällaisen riskin vaikuttavuus on korkea tai merkittävä, koska toteutuessaan vaikutukset koskisivat suurta määrää henkilötietoja ja ne muun muassa aiheuttaisivat pysyvän luottamuksen menetyksen rekisterinpitäjään. Riskin todennäköisyys on kuitenkin erittäin pieni, jos suhteutetaan Microsoftin pilvipalveluissa olevan datan määrä niissä tapahtuneisiin tietovuotoihin. (Bishop 2021.) Siksi sen riskiluku jää hyvin matalaksi. Muut tunnistetut riskit vaihtelivat vaikuttavuudeltaan matalasta korkeaan, mutta jokaisen riskin toteutuminen arvioitiin korkeintaan epätodennäköiseksi. Korkeimman riskiluvun synnytti käsittelytoiminnon tietokannan manuaaliseen päivittämiseen liittyvä tapahtumaketju, joka käsitellään seuraavassa luvussa.

4.2.2 Tasapainotesti

Sairauspoissaolotietojen käsittely tämän kehitystyön kontekstissa perustui rekisterinpitäjän oikeutettuun etuun. Oikeutettu etu tässä tapauksessa tarkoittaa sitä, että koska rekisteröity on rekisterinpitäjään työsuhteessa, on rekisterinpitäjällä ”oikeus” tulla raportoiduksi työntekijänsä poissaolosta. Oikeutus ei kuitenkaan ole automaattinen, ja se edellyttää että rekisteröidyn edut ja oikeudet huomioidaan erityisen tarkkaan. Oikeutettuun etuun liittyvät intressit voivat vaihdella merkityksettömistä hyvin tärkeisiin, ja vaikutukset rekisteröidyn oikeuksiin voivat vaihdella vähäisestä vakavaan. Työnantajan oikeus käsitellä työntekijän henkilötietoja perustuu myös useaan lakiin, kuten muun muassa työsopimus-, työaika- ja kirjanpitolakiin.

Tämän opinnäytetyön tapauksessa rekisterinpitäjällä oli jo valmiiksi pääsy haluttuun henkilötietoon eli sairauspoissaoloihin, mutta kyseisen tiedon uudenlainen käsittely saattoi aiheuttaa uudenlaisia riskejä rekisteröidyille. Rekisterinpitäjä onkin veloitettu punnitsemaan rekisteröidyn oikeuksia omia oikeuksiaan vasten aina kun uusi käsittelytoiminto otetaan käyttöön. Tämä tehdään niin kutsutulla tasapainotestillä. Tasapainotesti kattaa kuusi vaihetta, joista laaditun kirjallisen kuvauksen tarkoitus on osoittaa, että toiminta on tietosuojasetuksen mukaista ja päätöksenteko on läpinäkyvää (Tietosuojavaltuutetun toimisto 2022h.)

Koska tässä opinnäytetyössä kehitettyä raportointityökalua ei kuitenkaan otettu sellaiseen tuotantoon, on alla esitelty tasapainotesti ainoastaan alustava. Jotta tasapainotestiä

voisi todellisuudessa käyttää perustelemaan oikeutuksia, tulisi rekisteristä vastaavan henkilön tai tahon hyväksyä tehdyt valinnat. Tasapainotesti esitellään kuitenkin sellaisena, kuin se tässä tapauksessa saattaisi olla.

Ensimmäiseksi punnittiin käsittelyperusteen sopivuutta aiotulle käsittelytoiminnoille. Onko oikeutettu etu sopivampi käsittelyperuste kuin muut? Tässä tapauksessa mikään muu käsittelyperuste ei sopinut käsittelyn perusteeksi, joten oikeutettu etu oli ainut millä tasapainotestissä saatettiin edetä. Muiden käsittelyiden perusteet on esitelty kappaleessa 3.6.5.

Seuraavaksi arvioitiin ns. perusvaatimusten täyttymistä. Perusvaatimuksiin kuuluu, että etu on lainmukainen, etu on selkeästi ilmaistu, ja edun täytyy edustaa todellista ja välitöntä tarvetta. Spekulatiivinen etu ei käy. Käsittely todettiin lainmukaiseksi tietosuojalain 1050/2018 kanssa. Selkeästi ilmaistuna etu sairauspoissaolojen paremmasta raportoinnista on se, kun työnantaja saa laajan ja kattavan, mutta tarvittaessa myös tarkan kuvan sairauspoissaolojen aikamääreistä, työyksiköistä ja tarvittaessa sairastavista työntekijöistä, voi työnantaja suunnitella toimia, jotka vaikuttavat sairauspoissaolojen vähenemiseen. Tarve on todellinen ottaen huomioon, että aikaisemmat henkilöstöraportointijärjestelmät ovat poistuneet käytöstä järjestelmämuutosten myötä.

Kolmannessa vaiheessa arvioitiin, voisiko samaan lopputulokseen päästä keinoilla, joilla rekisteröidyn yksityisyyteen puututtaisiin vähemmän. Käsittelytoiminnon tarkkuus on tarvittaessa työntekijätasoinen, eli esihenkilö saa näkyviinsä alaisensa sairauspoissaolojen tilastitiikan. Koska työhyvinvointitoimenpiteiden kohdentaminen tapahtuu tarvittaessa yksilötasolla, pohdittiin että tarkkuus on sitä kautta perusteltua. Opinnäytetyöntekijä ehdotti, että koska tietolähteessä oli käytettävissä myös muita henkilön identifioimiseen sopivia sarakkeita kuin nimi tai henkilötunnus, tulisi sellaisen käyttöä harkita. Esimerkiksi työntekijänumero on tunnus, jonka ainoastaan työnantaja pystyy helposti yhdistämään henkilöön. Se on ainoastaan organisaation sisäisessä käytössä. Potentiaalisessa riskien toteutumistilanteessa, jossa käsittelytoiminnon tiedot vuotavat, olisikin tällöin työntekijän nimen ja sairauspoissaolotiedon sijaan vuotaneena työntekijänumero ja sairauspoissaolotieto, ja potentiaalinen riski rekisteröidyn oikeuksille olisi huomattavasti pienempi.

Näin saavutetulla riskien pienennyksellä olisi kuitenkin kääntöpuolensa. Käsittelytoiminnon käyttäjät eli esihenkilöt eivät tunnista työntekijänumeroista alaisiaan. Näin ollen esihenkilöt joutuisivat turvautumaan johonkin kolmanteen osapuoleen selvittääkseen kuka on tietyn työntekijänumeron takana. Tämä lisäisi työtä, mikä ei ole automatisoinnin tarkoitus, tai se voisi aiheuttaa tarpeen uusien henkilötietorekisterin synnylle. Esihenkilöllä voisi tätä varten olla esimerkiksi rekisterin ote, jossa olisi hänen alaisensa nimet ja työntekijänumerot. Näin

ollen oltaisiin päädytty liki samaan tilanteeseen kuin alussa, kuitenkin siten että olisi syntynyt uusia käsittelytoimintoja ja uusia riskejä.

Neljännessä vaiheessa arvioitiin, syrjäyttääkö rekisterinpitäjän etu todella rekisteröidyn edut ja oikeudet. Tietosuojavaltuutetun toimiston ohjeistuksena on lähestyä ongelmaa kysymyksillä kuten ”millaista hyötyä henkilötietojen käsittelystä olisi” ja ”millaista haittaa niiden käsittelemättä jättämisestä olisi”. Arvion mukaan käsittelytoiminto parantaisi työnantajan edellytyksiä kohdentaa työhyvinvointihankkeita, oli kyseessä sitten laajempaan kokonaisuuteen kuten työyksikköön- tai toimialueeseen liittyvä hanke tai yksittäistä työntekijää koskeva huomio. Edelleen tulisi kuitenkin pitää mielessä oikeutetun edun perusvaatimukset, jotka edellyttävät että etu ei ole spekulatiivinen. Arvion mukaan haittoiksi voitaisiin lukea esimerkiksi se, että kun työhyvinvointihankkeita ei osata kohdentaa, syntyy sairauspoissaoloista vaikutuksia toiminnan sujuvuuteen sekä organisaation talouteen.

Edellä mainittua hyötyä tulisi peilata konkreettisiin sekä spekulatiivisiin seurauksiin, joita käsittelystä voi rekisteröidylle olla. Niitä voivat olla esimerkiksi tulevat päätökset, toimet tai tilanteet, joissa käsittely voi johtaa henkilön syrjimiseen. Lisäksi neljännessä vaiheessa tulee vielä pohtia, että osaisiko rekisteröity odottaa, että hänen tietojaan käytetään aiotulla tavalla, tai olisiko todennäköistä, että rekisteröity vastustaisi tietojensa käsittelyä? Opinnäytetyöntekijän arvion mukaan työntekijällä on perusteltu syy odottaa, että hänen esihenkilönsä on tietoinen hänen töistä poissaolostaan, eikä hänellä olisi syytä vastustaa hänen tietojensa käsittelyä.

Näiden edellä mainitun neljän vaiheen jälkeen rekisterinpitäjä on jo huomattavasti valmiimpi arvioimaan oman etunsa painoarvoa suhteessa rekisteröidyn etuihin ja oikeuksiin. Lisäarviointia tarvitaan kuitenkin, mikäli on epäselvää mihin suuntaan tasapaino kallistuu. Jos tasapainoa halutaan vielä ”kallistaa” oikeutetun edun puolelle, voidaan kokeilla ottaa käyttöön tietosuojan lisätakeita. Tasapainotestin lopputuloshan riippuu kokonaisarvioista. Mitä huomattavampi vaikutus käsittelyllä on rekisteröidyn kannalta, sitä enemmän rekisterinpitäjältä edellytetään asianmukaisia takeita tietosuojalle. Rekisterinpitäjä voi esimerkiksi eriyttää toimintojaan teknisillä ja organisatorisilla toimenpiteillä, jotta käsittelytoiminto olisi suojatumpi. Laajat anonymisointi- tai salaustekniikat voivat myös parantaa tilannetta.

Tämän opinnäytetyön kehitysosassa kehitetylle uudelle käsittelytoiminnolle tehtyä tasapainotestiä ei kuitenkaan viety päätökseensä asti, koska käsittelytoimintoa ei otettu sellaiseen tuotantoon. Opinnäytetyösuunnitelmasta poiketen kehitystyön tietolähteenä ei käytetty tuotannossa olevaa henkilöstöresurssin tietokantaa tai -kantoja. Tietokantaratkaisun toimittanut ohjelmistoyhtiö ei ollut halukas tämän opinnäytetyön aikaikkunassa tarjoamaan rajapintaa, jonka päälle kehitystyötä olisi voitu alkaa tehdä. Tietokantayhteyden sijaan

raportoinnin kehitystyössä käytettiin henkilöstöhallinnon käyttöliittymän avulla tietokannoista vietyjä koontitiedostoja, jotka kattoivat opinnäytetyösuunnitelmassa määritellyt aiheet sekä aikamääreet tarpeeksi pitkältä ajanjaksolta, jotta visualisoitujen tietojen oikeellisuutta oli mahdollista arvioida.

Koontitiedostojen vienti tehtiin tuotantokannasta web-käyttöliittymän kautta tuoduilla tiedoilla (CSV), jotka kattoivat halutun aikavälin tiedot. Tämän prosessin toistaminen uuden ”raportointikannan” päivittämiseksi olisi aiheuttanut sen, että tietojen kopioita syntyisi usealle eri työasemalle. Mikäli operaatio toistettaisiin säännöllisesti, jokainen kerta mahdollistaisi uusien kopioiden syntymisen, ja lopulta kopio saattaisi päätyä joko väärään kansioon, väärällä tavalla siirretyksi, tai unohtua väärään paikkaan jolloin sen tietosuoja olisi uhattuna. Monta iteraatiota manuaalista toimintaa ihmisen tekemänä voi ainoastaan nostaa virheiden tapahtumisen mahdollisuutta, ei laskea.

Tietokantayhteyden puuttuminen vaikutti kehitystyön tietosuojan ja tietoturvan toteutukseen. Manuaalinen työ tietojen päivittämisessä todettiin erikseen vielä tietosuojan vaikutustenarvioinnissa korkeariskiseksi toiminnaksi, mikä ennestään nosti kynnystä ottaa raportit sellaisenaan käyttöön.

Mikäli käsittelytoiminto olisi otettu käyttöön, tulisi tämänkaltainen päätöksenteko ja päätökseen johtanut pohdinta dokumentoida, jotta rekisterinpitäjä voisi tarvittaessa osoittaa toimintansa olevan tietosuoja-asetuksen mukaista. Avoimuus ja läpinäkyvyys tukevat oikeutetun edun käyttöä käsittelyperusteena, joten rekisterinpitäjän tulisi olla kykenevä perustelemaan ratkaisunsa rekisteröidylle.

4.2.3 Aineiston tietoturvallinen käsittely kehitysvaiheessa

Kehitystyössä käytetty aineisto vietiin kahdeksi eri tiedostoksi henkilöstöhallinnon tietojärjestelmästä. Toinen kattoi sairauspoissaolojen tiedot, ja toinen osaamisen kehittämisen tiedot. Koska heti ensimmäisestä viennistä alkaen tiedostoista alkoi muodostua kopioita, alettiin tiedostoille tehtyjä toimenpiteitä seurata dokumentaatiolla. Dokumentaation merkitys korostui etenkin kehitystyön päätyttyä, jolloin aineistot kokonaisuudessaan tuhottiin.

Ensimmäiset kopiot syntyivät viennissä käytetylle työasemalle, ja seuraava heti tiedostot salatulle muistitikulle kopioitaessa. Varmuuskopiot alkuperäisestä aineistosta luotiin niin ikään tässä vaiheessa. Alkuperäinen, todelliset henkilötiedot sisältävä aineisto ei siirtynyt enää tämän jälkeen salatuilta muistitikuilta muualle.

Muistitikkuja käytettiin etenkin siksi, että pandemiatilanteen vuoksi kehitystyöhön osallistuneista työntekijöistä osa oli etätöissä, ja aineisto olisi pitänyt siirtää internetin yli. Lisäksi

aineiston lopullinen tuhoaminen on helpompaa muistitikulta kuin kovalevyiltä, joka edellyttäisi useita päällekirjoituskertoja.

Muistitikut salattiin Microsoft BitLocker -ohjelmistolla, jolla myös kehitystyöhön osallistuneiden työntekijöiden kovalevyillä olevat käyttäjäprofiilit tiedostoineen on salattu. Muistitikujen salausalgoritmi oli AES ja palautusavaimet säilytettiin ainoastaan yksinkertaisina paperitukosteina.

Kehitystyössä käytetty anonymisoitu aineisto sen sijaan tallennettiin opinnäytetyön tekijän työasemalle ja siitä edelleen lopulta myös Power BI-palveluun. Aineiston kopioiden seurannan dokumentaatiota jatkettiin myös anonymisoidun aineiston osalta, koska aineisto sisälsi yksityiskohtaisia tietoja REDUn yksiköiden sairauspoissaolojen määristä ja laaduista. Näin ollen nekin kuuluivat tuhottaviin aineistoihin kehitystyön päätyttyä.

4.2.4 Aineiston anonymisointi ja normalisointi

Anonymisointi toteutettiin Microsoft Excel -ohjelmistolla. Ensin luotiin vastaava määrä satunnaisia kuuden kirjaimen arvoja, kuin oli eri henkilöiden nimiä. Satunnaisten arvojen luonnissa käytettiin TEXTJOIN, INDEX ja RANDARRAY -funktioita:

=TEXTJOIN("",1,INDEX(aakkoset,RANDARRAY(6,1,1,26,TRUE)))

jossa "aakkoset" on nimetty taulukon osa sisältäen kirjaimet A-Z, joista RANDARRAY palautti satunnaiset kuusi kirjainta em. funktion sisältävään soluun. Kun oli tiedossa, että eri nimisiä henkilöitä oli esimerkinomaisesti 200 kappaletta, luotiin siis ensimmäiseksi 200 satunnaista kuuden kirjaimen merkkijonoa. Satunnaiset merkkijonot ja henkilöiden nimet kopioitiin vierekkäin aputaulukkoon. Tämän jälkeen aineistossa olevat nimet korvattiin aputaulukon ja VLOOKUP -funktion avulla satunnaisilla merkkijonoilla. Lopuksi aputaulukko tuhoettiin.

Sekä sairauspoissaolojen että osaamisen kehittämisen aineistoja tarvitsi myös siistiä. Asiaan liittymätöntä informaatiota suodatettiin, duplikaatti- ja puutteellisia rivejä poistettiin sekä ylätunnisteita selkeytettiin. Lisäksi osaamisen kehittämisen aineistossa työntekijän koulutuksesta oli tiedossa aloitus- ja lopetuspäivämäärä sekä kesto yhteensä tunteina. Aineistoon kohdistettiin PowerQuery-muunnos, joka jakoi aloitus- ja lopetuspäivämäärän sekä niiden välisen ajan siten, että yhdellä rivillä oli yksi päivä, ja koulutuksen kesto tunteina jaettiin näille riveille tasan. Tämä ei välttämättä kuvastanut todellisuutta täysin oikein, sillä esimerkiksi maanantaina ja tiistaina yhteensä kymmenen tuntia kestänyt koulutus oli saattanut olla maanantaina täysi työpäivä, ja tiistaina puolikas päivä. Muunnoksen jälkeen maanantain ja

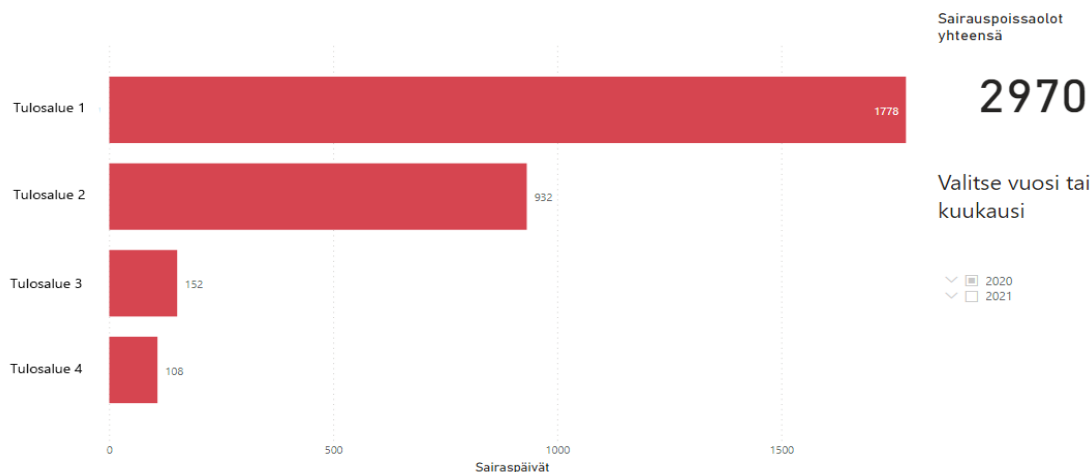
tiistain kesto oli kuitenkin viisi tuntia. Useamman päivän mittakaavassa tieto oli kuitenkin oikein, eikä sitä tarkemmaksi ollut mahdollista saada aineiston rajoittuneisuudesta johtuen.

4.3 Iteraatiot anonymisoiduilla tiedoilla

4.3.1 Ensimmäinen iteraatio

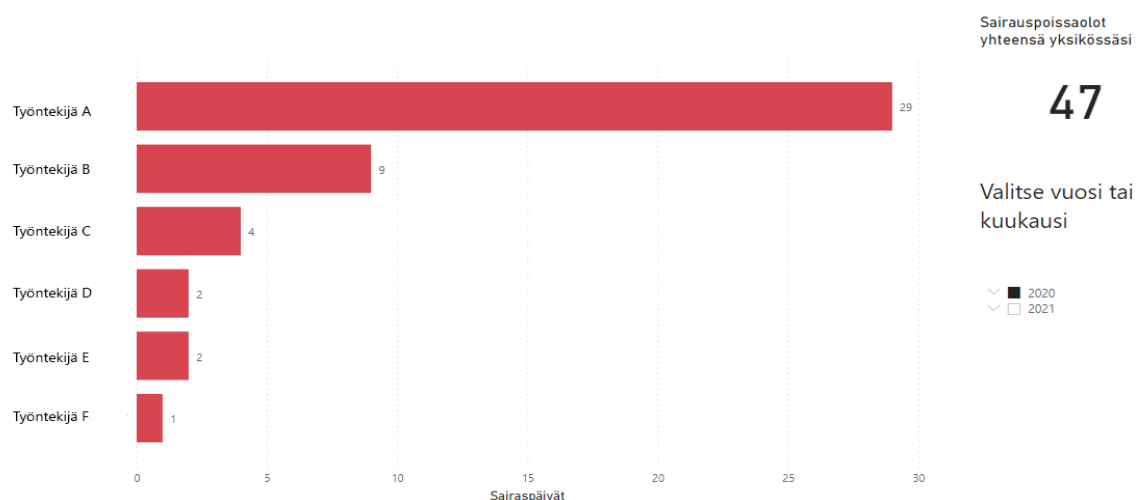
Sairauspoissaolojen tietolähde sisälsi 16 saraketta tietoa ja osaamisen kehittämisen tietolähde viisi saraketta. Kuten kehitystyöryhmän kanssa oli aloituspalaverissa sovittu, raporttien ensimmäiseen versioon pyrittiin visualisoimaan mahdollisimman monipuolisesti saatavilla olevaa tietoa, joista voisi sitten jatkaa kehitystyötä erilaisia toteutuksia karsien ja hyödyllisiä parantaen. Sairauspoissaolojen raportin ensimmäiseen versioon tehtiin opinnäytetyön tekijän toimesta seitsemän eri visualisointia, jotka esiteltiin REDUn osaamisen kehittämisen työryhmälle. Ensimmäinen visualisoinneista esitteli sairauspoissaolojen määrät pylväsdiagrammina. Pylväsdiagrammin ylätaso oli jaoteltu tulosalueittain (Kuva 2), ja näkymä oli porattavissa yksiköihin (eng. "drill down") REDUn konserni-organisaatiokaavion sekä ammatillisen koulutuksen toimialarakenteen mukaisesti (REDU 2022a). Yksikkö oli edelleen porattavissa työntekijätasolle asti (Kuva 3). Ajatuksena oli, että toimiala- sekä yksikötasojen näkymät olisivat samat kaikille esihenkilöille, mutta työntekijätasolle porattaessa esihenkilö näkisi ainoastaan omien alaistensa tiedot.

Sairauspoissaolot tulos- ja toimialoittain, kalenteripäiviä



Kuva 2. Ylätason näkymä

Sairauspoissaolot henkilöittäin, kalenteripäiviä



Kuva 3. Esihenkilön näkymä omaan tulosalueeseen tai yksikköön porattaessa ("Drill-down")

Tämän toteutus rivitason suojauksella (RLS) sekä käyttöoikeuksilla ei ollut vielä tässä vaiheessa valmis, ja toteutustapaa jouduttiinkin muuttamaan seuraavassa iteraatiossa. Toimi-alue- ja yksikötason näkymä sen sijaan osoittautui toivotuksi ominaisuudeksi, ja se päättyi-kin viimeiseen versioon asti. Näkymää oli myös mahdollisuus suodattaa aikaperusteisesti vuoden ja kuukauden mukaan.

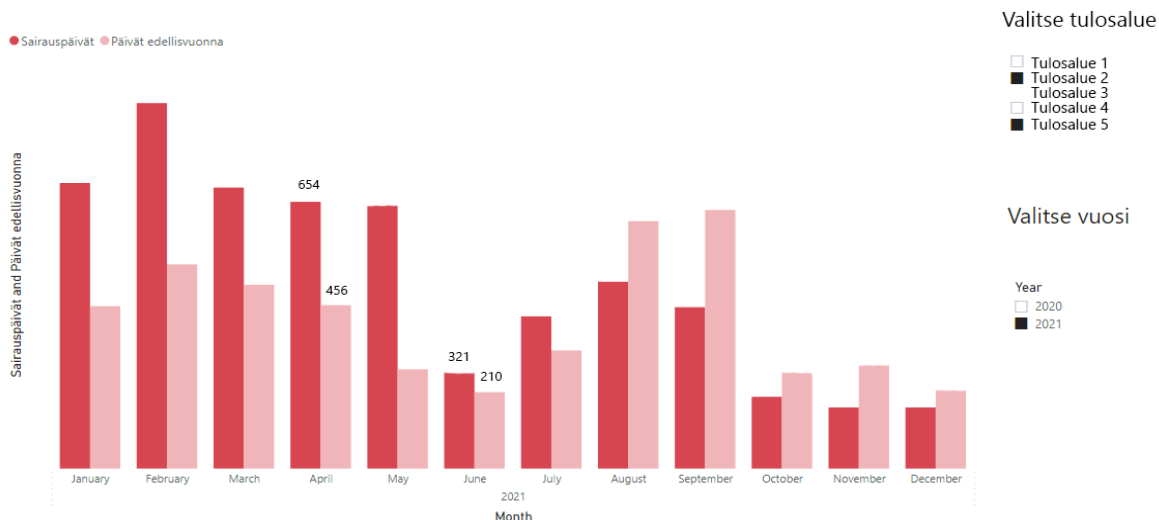
Toinen visualisointi ilmensi sairauspoissaolojen koodeja. Pylväsdiagrammin jaottelu ja suodatus oli identtinen ensimmäisen visualisoinnin kanssa, sillä muutoksella että pylväät olivat "kasattuja" sairauspoissaolojen koodeilla. Koodit eivät sisältäneet sairauspoissaolojen syitä, vaan ilmensivät poissaolon vaikutusta palkanmaksuun.

Kolmas visualisointi kuvasi näiden koodien suhteellista jakautumista tulosalueiden, yksiköiden ja työntekijöiden sisällä. Jokainen palkki oli siis 100% ja sen sisältö oli jaoteltu koodien suhteellisilla määrillä suhteessa tulosalueen tai yksikön kokoon. Tämän oli tarkoitus helpottaa koodien jakauman vertailua tulosalueiden ja yksiköiden kesken johtuen niiden isoista kokoeroista. Lisäksi visualisointi olisi tarjonnut mahdollisuuden vertailla koodien jakaumaa eri työntekijöiden kesken. Työryhmän kanssa visualisointiin lähemmin perehdyttäessä tiedon luotettavuudesta löydettiin kuitenkin puutteita. Tietolähteeseen päätyvien sairauspoissaolojen koodeihin vaikuttivat useat eri tekijät, kuten sairauspoissaolon jatkuminen työntekijän omasta ilmoituksesta lääkärin määräämään, ja työntekijän oma toiminta näistä muutoksista raportoidessa. Sairauspoissaolojen koodien visualisointi jätettiin siksi jatkokehityksestä pois.

Neljäs visualisointi kuvasi kuukausittaisia sairauspoissaoloja vuoden ajalta histogrammina. Edellisen vuoden vastaava kuukausi oli haaleampana palkkina "taustalla" jotta vertailu eri

vuosien välillä olisi helppoa. Näkymä oli suodatettavissa ajan, toimialueiden ja yksiköiden perusteella. Myös tämä visualisointi (Kuva 4) säilyi viimeiseen versioon asti.

Sairauspoissaolojen vuosivertailu



Kuva 4. Vuosivertailu edelliseen vuoteen

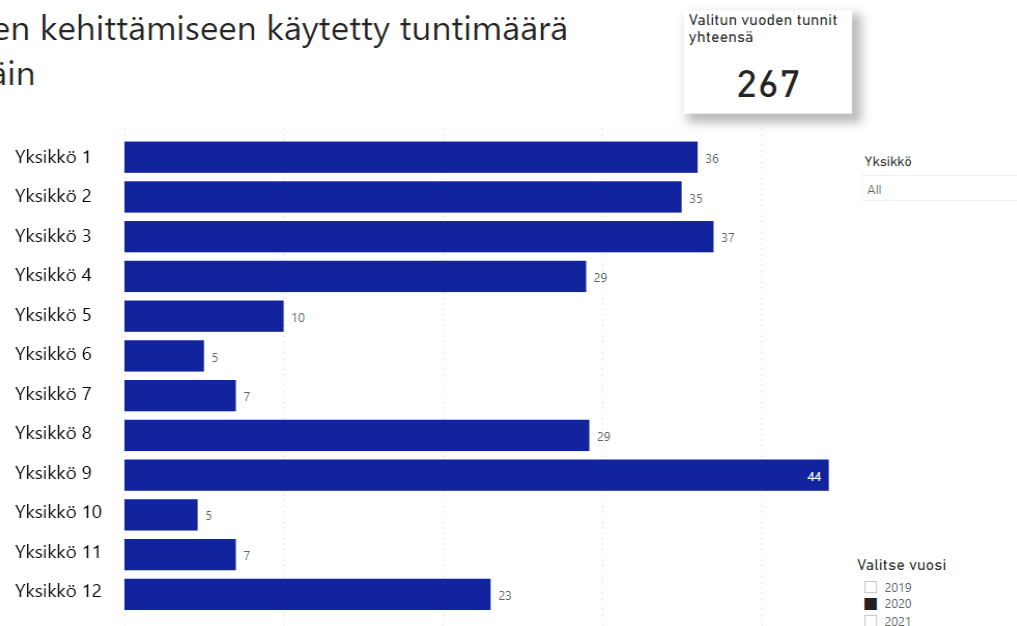
Työryhmän kanssa visualisointiin lähemmin perehdyttäessä siinä havaittiin kuitenkin ongelma datan jakautumisessa eri kuukausille. Mikäli työntekijällä alkoi esimerkiksi kahden kuukauden sairauspoissaolo tammikuun alussa, ei tämä poissaolo jakaantunut tasaisesti tammi- ja helmikuulle, vaan nosti tammikuun arvoa myös helmikuun poissaolon määrällä. Tähän ongelmaan pureuduttiin seuraavassa iteraatiossa.

Viides visualisointi kuvasi sairauspoissaolojen aloituspäivän jakautumista eri viikonpäiville työntekijöittäin. Visualisointia oli mahdollista suodattaa ajan, toimialueiden, yksiköiden ja sairauspoissaolon pituuden perusteella. Näkymästä olisi ollut mahdollista erottaa mikäli esimerkiksi jokin tietty viikonpäivä olisi yliedustettuna lyhyiden poissaolojen aloituspäivänä. Visualisointi ei kuitenkaan tarjonnut näkyviä trendejä minkä lisäksi sen toteutus seuraavassa iteraatiossa rivitason suojauksen kanssa olisi tuonut haasteita. Tämä visualisointi jätettiin jatkokehityksestä pois.

Kuudes visualisointi kuvasi sairauspoissaolojen koodien jakautumista eri viikonpäiville, eli alkoivatko esimerkiksi esihenkilön luvalla tapahtuvat sairauspoissaolot tyypillisemmin jonnain tiettyinä päivinä. Näkymä oli suodatettavissa kuten edellisetkin. Myös seitsemäs visualisointi kuvasi sairauspoissaolojen koodeja, pyrkien esittämään niiden vuotuista vaihtelua. Näiden visualisointien jatkokehityksestä luovuttiin kuitenkin samasta syystä kuin muidenkin sairauspoissaolojen koodien esityksistä.

Sairauspoissaolojen lisäksi ensimmäiseen iteraatioon kehitettiin kuusi erilaista visualisointia osaamisen kehittämisen resurssin käyttöä kuvaamaan. Näiden visualisointien käyttöoikeuksien idea oli sama kuin sairauspoissaoloissa, eli yksiköt olivat porattavissa työntekijöiksi, mikäli raporttia käyttävä esihenkilö poraisi omaan yksikköönsä. Ensimmäinen kuvasi osaamisen kehittämiseen käytetyn tuntimäärän toteumaa yksiköittäin pylväsdiagrammissa (Kuva 6).

Osaamisen kehittämiseen käytetty tuntimäärä yksiköittäin



Kuva 6. Tuntimäärän toteuma pylväsdiagrammissa

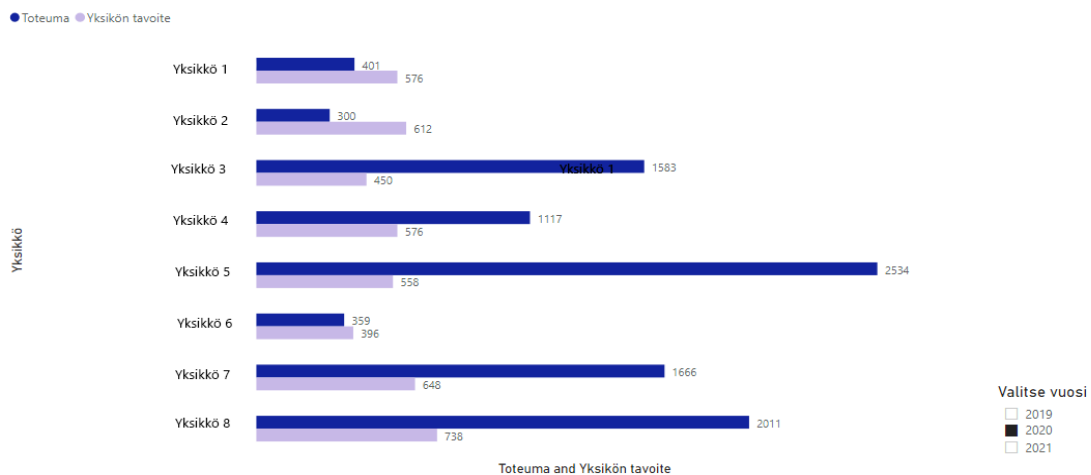
Yksiköt ovat kuitenkin varsin eri kokoisia, joten ensimmäisen visualisoinnin tueksi kehitettiin visualisointi, joka esitti toteuman suhteessa yksikön tavoitteeseen (Kuva 7). Yksikön tavoite esitettiin toteuman viereisenä pylväänä pylväsdiagrammissa, ja se laskettiin kertomalla tavoite (18h) yksikön kokonaishenkilömäärällä. Ajatuksena oli, että esihenkilö pystyisi näin seuraamaan koko yksikkönsä menestymistä tavoitteen saavuttamisessa. Koska laskenta oli riippuvainen käyttäjän tekemistä valinnoista, kuten korostus-, vuosi- tai yksikkövalinnoista, tuli laskennan muuttua näiden valintojen mukana:

$$\text{Yksikön_tavoite} = [\text{Tavoite}] * \text{SUM}(\text{'Henkilöstö' [Hlö_lkm]})$$

Ylläoleva DAX-syntaksi palauttaa arvon mittarille, jonka nimi on "Yksikön_tavoite". Yksikön tavoite on alla olevassa kuvassa vaaleampi palkki, johon toteumaa verrataan. Arvo lasketaan siten, että henkilöstötaulun henkilölukumäärä-sarakkeen summa, johon käyttäjän tekemät valinnat ja suodatuksat vaikuttavat, kerrotaan ennalta määrätyllä muuttujalla "Tavoite", joka siis oli 18 tuntia.

18h vuositavoitteen toteuma: Yksikön toteuma suhteessa yksikön tavoitteeseen

Yksikön tavoite = Yksikön henkilömäärä * 18h



Kuva 7. Yksikön toteuma suhteessa tavoitteeseen

Tämä sama asia esitettiin myös suhteellisesti, eli sataprosenttisina pylväinä, jotka oli jaettu toteumaan sekä puuttuvaan osaan tavoitteesta prosentuaalisesti esitettynä (Kuva 8). 100 % pylväsdiaagrammi koostuu yhdestä tai useammasta arvosta, joiden suhteelliset osuudet arvojen yhteenlasketusta summasta esitetään eri kokoisina- ja värisinä paloina pylväässä. Tällä visualisoinnilla häivytettiin yksiköiden koon merkitys niiden keskinäisessä vertailussa. Ajatuksena oli, että yksikköä, jossa oli 10 työntekijää, voisi vertailla tavoitteen saavuttamisessa yksikköön, jossa oli 100 työntekijää. Tämä toteutettiin kohdistamalla laskentaan conditionaalinen ehto, joka toteutettiin DAX:n syntaksin IF-lausekkeella, jossa hyödynnettiin jo edellisessä visualisoinnissa käytettyä "Yksikön tavoite"-muuttujaa:

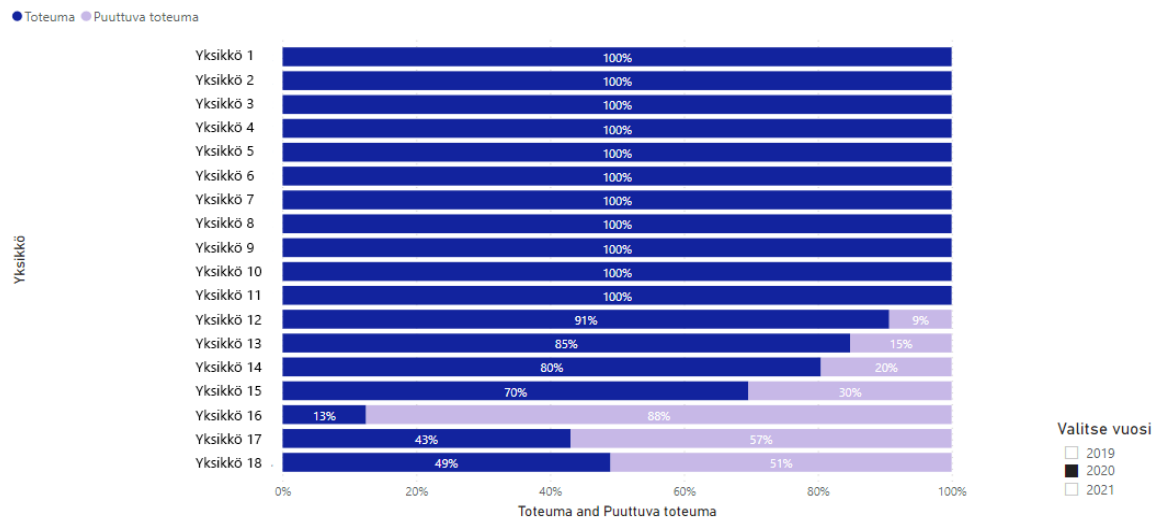
IF('Henkilöstö'[Yksikön_tavoite minus toteuma] > 0,

[Yksikön_tavoite] - SUM('Henkilöstö'[toteuma]))

Jossa jos yksikön tavoitteen ja toteuman erotus on enemmän kuin nolla, niin palautetaan arvo, joka on yksikön tavoitteen ja yksikön toteuman erotus, eli mikäli yksikkö ei ole vielä saavuttanut tavoitettaan, niin puuttuva osa lasketaan ja lisätään pylvääseen omalla suhteellisella osuudellaan.

18h vuositavoitteen toteuma yksiköittäin, suhteellinen

Koko palkki (100%) = Yksikön tavoite eli 18h * yksikön henkilömäärä
 Vaaleansininen = Toteutunut tavoite
 Tummansininen = Puuttuva osa tavoitteesta



Kuva 8. Toteuman suhteellinen vertailu

Neljäs visualisointi kuvasi tavoitteen toteutumista henkilöittäin. Käyttöoikeuden idea oli sama kuin sairauspoissaoloissa, eli esihenkilö näkisi ainoastaan omien alaistensa tiedot. Tavoite ja toteuma esitettiin viereisinä pylväinä diagrammissa kuten yksiköidenkin toteuman ja tavoitteen visualisoinnissa.

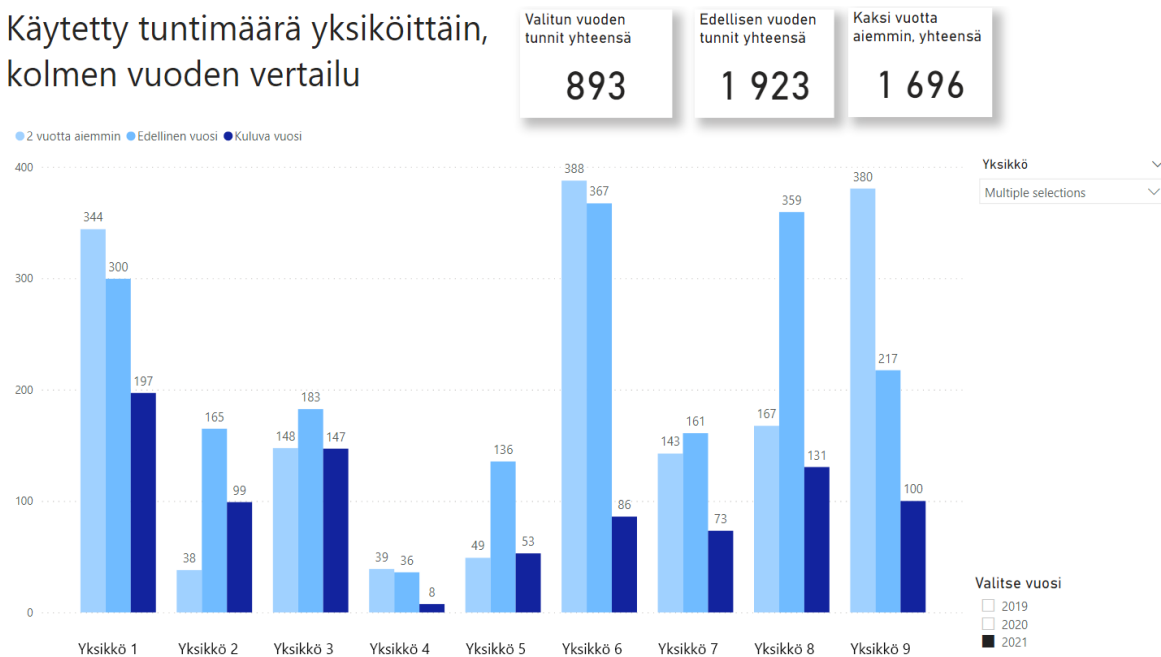
Viides visualisointi kuvasi käytettyä tuntimäärää yksiköittäin kolmen vuoden vertailuna (Kuva 9). Pylväsdiaagrammilla oli siis mahdollista nähdä haluttujen yksiköiden käytetyn tuntimäärän kehitys kolmen vuoden jaksolla. Käyttäjän valintojen mukaan muuttuvan ajan kanssa tapahtuvaan laskentaan on DAX:ssa useita eri funktioita. Valitun ajankohdan vertailu vuotta aiempaan on helppoa "SAMEPERIODLASTYEAR"-nimisellä funktiolla:

Kesto edellisvuonna = CALCULATE([Total Kesto], SAMEPERIODLASTYEAR(Date_Table[DateKey]))

jossa "Total kesto"-niminen mittari ensin laskee käyttäjän valitseman ajanjakson tunnit, mutta arvoksi palautetaankin edellisen vuoden arvot samalta ajanjaksolta. Kaksi vuotta aiemmin tapahtunutta varten joudutaan kuitenkin käyttämään eri funktiota, joskin "parallel-period"-funktioilla olisi voinut palauttaa myös edellisen vuoden arvot muuttamalla arvon "-2" arvoksi "-1":

Kesto kaksi vuotta aiemmin = calculate([Total Kesto], parallelperiod(Date_Table[DateKey], -2, year))

Käytetty tuntimäärä yksiköittäin, kolmen vuoden vertailu



Kuva 9. Yksiköiden toteuman kolmen vuoden vertailu

Kuudes visualisointi kuvasi koulutuksiin osallistuneiden ja osallistumattomien jakaumaa jaettuna pylväskuviona. Jokainen yksikkö oli esitetty omana sataprosenttisena pylväänsä, jonka sisältö oli jaettu kahteen osaan eli heihin, joilla oli vähintään yksi tunti osaamisen kehittämisen resurssia käytettynä ja heihin, joilla ei ollut. Tietolähde itsessään ei sisältänyt tietoa koulutuksiin osallistumattomista työntekijöistä, mutta heidän määränsä oli mahdollista laskea vähentämällä yksikön kokonaishenkilömäärästä koulutuksiin osallistuneiden määrä. Osa- ja kokoaikaisten työntekijöiden määrää ei kuitenkaan pystytty tässä vaiheessa erottelemaan tarpeeksi lyhyellä aikavälillä, joten tämä visualisointi jäi pois seuraavista iteraatioista.

4.3.2 Toinen iteraatio

Ensimmäisessä iteraatiossa ilmeni ongelma tietojen porattavuuden suhteen, joka koski niin sairauspoissaolojen kuin osaamisen kehittämisen raporttien toteutusta. Tarkoituksena oli, että yleistason näkymät, jotka esittelivät tulosalueita ja yksiköitä, näkyisivät määrällisine arvoineen kaikille käyttäjille eli REDUn esihenkilöille. Sen sijaan henkilötason datan olisi pitänyt suodattua esihenkilön ja työntekijän suhteen perusteella, eli sen perusteella, oliko kyseinen esihenkilö tarkastelun kohteena olevan yksikön esihenkilö vai ei. Käytännössä tämän oli tarkoitus ilmetä siten, että mikäli esihenkilö porautuisi ensin tulosalueen sisälle, josta hän edelleen porautuisi tiettyyn yksikköön, hän näkisi yksikön sisällön työntekijöinä, mikäli nämä olisivat hänen alaisiaan. Mikäli eivät olisi, eli jos hän porautuisi jonkun toisen esihenkilön yksikköön, hän ei näkisi työntekijöiden tietoja.

Ensimmäisessä iteraatiossa tieto tulosalueen, yksikön ja työntekijän tiedoista olivat samassa taulussa. Tai hieman kääntäen ilmaistuna, taulu sisälsi työntekijän tiedot, ja työntekijöiden tiedot sisälsivät määrällisten tietojen, kuten sairauspoissaolojen tai osaamisen kehittämisen resurssin lisäksi työntekijän tulosalueen ja yksikön. Näin ollen tulosalueiden ja yksiköiden visualisoinnit, jotka käytännössä olivat erilaisia summapylväitä aikamääreillä, laskettiin ”työntekijätaulusta”. Rivitasonsuojaus (RLS) kohdistetaan jonkin taulun sarakkeen sisällön perusteella. Eli jos taulu sisältää esimerkiksi tiedot kuten kuvassa 10:

Työntekijä ▼	Yksikkö ▼	Tulosalue ▼
Essi Esimerkki	Yksikkö 1	Tulosalue A
Mikko Mallikas	Yksikkö 2	Tulosalue B
Erkki Erikoinen	Yksikkö 4	Tulosalue B

Kuva 10: Esimerkkitaulu

ja rivitason suojauksen DAX ilmaisu on

[Yksikkö] = ”Yksikkö 1”

näkisi tälle DAX ilmaisulle roolitettu käyttäjä taulusta ainoastaan yksikön yksi tiedot, eli ensimmäisen rivin tiedot, ja ainoastaan niitä tietoja käytettäisiin laskennoissa tai visualisoinneissa. Rivitason suojaus koskee kuitenkin koko tietojoukkoa, eli se antaa käyttäjän nähdä tai käyttää tietolähteestä ainoastaan ne rivit, joista edellä mainittu DAX ilmaisu palauttaa arvon TRUE (tosi). Tästä aiheutui se, että myös ne visualisoinnit, jotka koskivat yksikkö- ja tulosaluenäkymiä, suodattuivat käyttäjän roolituksen mukaisesti, ja he näkivät summamäärisissä pylväissä ainoastaan oman yksikkönsä summat.

Ongelma pyrittiin kiertämään jakamalla eri visualisoinneissa käytetty tieto eri tauluihin, sillä ajatuksella, että tarkan henkilötason tiedon sisältävä taulu oli irrallisena yksikkö- ja tulosaluiden tiedot sisältävästä taulusta. Tämä yksistään ei välttämättä ole riittävä ratkaisu, koska rivitason suojaus (RLS) koskee todellakin koko tietojoukkoa, eikä yksittäisiä tauluja. Jos rivitason suojaus määritellään tietojoukon taulussa X, se pyrkii ”leviämään” sieltä myös muihin tauluihin mikäli taulujen väliset suhteet sen mahdollistavat. Kardinaliteettiin voidaan kuitenkin vaikuttaa ja suojaussuodattimia käyttää joko molempiin suuntiin tai vain toiseen.

Rivitason suojauksen toteutusta kehitettäessä tuli selväksi, että siihen liittyi myös aspekteja jotka vaikuttivat suojauksen tietosuojan tasoon. Väärinymmärtäminen tai vääränlainen toteutus saattaisi aiheuttaa riskin henkilötietojen näkymiseen tahoille, joilla ei tietoihin pääsyä muuten olisi. Riskin vaikuttavuutta pienensi kuitenkin huomattavasti se, että pääsy raporttiin ja tietojoukkoon ylipäättään olisi hyvin rajattu, eli pääsy olisi vain REDUn esihenkilöillä. Mikäli rivitason suojauksessa olisi virheitä tai se ei toimisi toivotunlaisesti, riski rajoittuisi siihen

että ”väärän” yksikön esihenkilö näkisi jonkin toisen yksikön työntekijöiden tietoja. Jos rivitason suojauksen kaikkia teknisiä ominaisuuksia ei täysin ymmärretä, siihen lienee parasta suhtautua samaan tapaan kuin raporttien suodattimiin.

Myös kalenteripäivien jakautumisessa ajallisesti huomattiin virheitä ensimmäisessä iteraatiossa. Poissaolon keston liittyen kirjattu aikamäärä oli kirjattu jollekin tietylle aikavälille. Eli jos työntekijä esimerkiksi oli ollut kuukauden sairauslomalla ja tehnyt tästä ilmoituksen REDUn tietojärjestelmässä, aineistoon tallentui poissaolon alkamis- ja päättymispäivämäärä sekä kesto. Tämän visualisointi Power BI:ssä tarkoittaa kuitenkin sitä, että jompaankumpaan päivämäärään pitäisi sitoa esimerkin keston arvo eli kolmekymmentä kalenteripäivää. Mikäli se sidotaan alkamispäivämäärään, ja alkamispäivämäärä on jotain muuta kuin kuukauden ensimmäinen päivämäärä, keston pitäisi jatkua seuraavalle kuukaudelle. Visualisoinnissa alkamispäivän kuukausi saa silti mainitun arvon kokonaisuudessaan. Päättymispäivämäärään sitominen olisi myös aina väärin kohdistunut. Tämän vuoksi aineistossa olleisiin kestoihin kohdistettiin PowerQuery -muunnos, jonka lopputuloksena yhdellä rivillä oli aina sairauspoissaolon yksi päivä ja päivällä oli kalenteriin sidottu arvo.

4.3.3 Kolmas ja viimeinen iteraatio

Kolmanteen iteraatioon karsittiin edelleen raportteja kehitystyöryhmän keskusteluiden pohjalta. Sairauspoissaolojen raportointiin jäi kolme eri visualisointia: sairauspoissaolot tulos- ja toimialueittain, sairauspoissaolot henkilöittäin sekä vuosivertailu, jolla saattoi vertailla eri yksiköiden sairauspoissaolojen määrää edellisiin vuosiin. Osaamisen kehittämisen raporttiin jäi viisi eri visualisointia: Osaamisen kehittämiseen käytetty tuntimäärä yksiköittäin, 18h vuositavoitteen toteuma henkilöittäin, yksikön toteuma suhteessa yksikön tavoitteeseen, yksikön toteuma suhteessa yksikön tavoitteeseen suhteellisesti esitettynä sekä vuosivertailu, jolla saattoi vertailla eri yksiköiden käytetyn tuntimäärän kehitystä kolmen vuoden jaksolla. Näistä jäljellejääneistä visualisoinneista on kuvat kappaleessa 4.3.1.

4.4 Kehitystyön päätös

Joulukuussa 2021 kehitystyöryhmän kanssa keskusteltaessa tultiin siihen tulokseen, että raportit visualisointineen olivat esittelykelpoisia REDUn esihenkilöille. Raporttien kehittäminen opinnäytetyön puitteissa päätettiin siis kolmanteen iteraatioon. Kehitystyön tässä vaiheessa ei edelleenkään oltu saatu yhteistyötä käyntiin tuotantokannan toimittajan kanssa rajapinnan avaamiseksi. Raporttien päivittäminen olisi vaatinut manuaalista tietojenkäsittelyä, joka todettiin tietosuojariskiksi tietosuojan vaikutustenarvioinnissa. Tästä syystä raportteja ei tämän opinnäytetyön puitteissa otettu tuotantoon.

Kehitystyön konstruktio oli siis kaksi erillistä Power BI -raporttia, joista toinen käsitteli sairauspoissaoloja ja toinen osaamisen kehittämiseen varatun resurssin toteutumista. Konstruktio esiteltiin REDUn pääallikkokokouksessa helmikuussa 2022 läsnä olleille esihenkilöille. Ennen toiminnallisuuden esittelyä esihenkilöitä kehoitettiin miettimään esittelyn aikana, olisiko heille konstruktiosta hyötyä heidän työssään. Toiminnallisuuden esittelyn jälkeen esihenkilöillä oli mahdollisuus esittää kommentteja ja kehitysideoita.

Kritiikkiä kohdistettiin sairauspoissaolojen raportoinnin tarkkuuteen. Yhden kommentin mukaan pelkkä numeerinen tieto sairauspoissaoloista ei välttämättä palvele tarkoitusta, olkoonkin että sen pystyisi kohdentamaan kuinka tarkasti hyvänsä. Kommentoijan mukaan tarvittaisiin enemmän tietoa sairauspoissaolojen syistä, jotta esihenkilöiden olisi mahdollista reagoida niihin oikein. Työnantajalla on rajatuissa tapauksissa oikeus käsitellä työntekijän terveydentilatietoja, mutta sillä ei kuitenkaan ole oikeutta kerätä tätä tietoa, eikä sitä siten voi olettaa käytettävän raportoinnissa (Työsuojeluhallinnon verkkopalvelu 2021). Rekisteriä sairaustiedoista saa pitää vain terveydenhuollon toimija. Yleisellä tasolla sairauspoissaolojen raporttia kuitenkin kiiteltiin selkeydestä sekä sen vähäisen informaation kuvaamisesta, mitä saatavilla oli ollut.

Osaamisen kehittämisen resurssin raportti keräsi eniten positiivisia kommentteja, joiden mukaan siitä olisi hyötyä esihenkilötyössä. Erään kommentin mukaan kyseisestä resursista maksetaan palkkaa kaikille, joten sen tasapuolisen käytön seurantaan tarvitaan toimivia työkaluja. Toisen mukaan resursseja tulisi ilman muuta seurata, jotta työnantaja tietäisi mitä työntekijä on hänelle annetuilla tunneilla tehnyt. Kolmannen mukaan jotain esittelyn kaltaista on toivottu jo pitkään. Muutama kommentti koski koko esitystä yleisellä tasolla ja sisälsi kiitettävää palautetta ja työkaluja toivottiin tuotantokäyttöön. Yhden kommentin mukaan esitys oli turhan nopea.

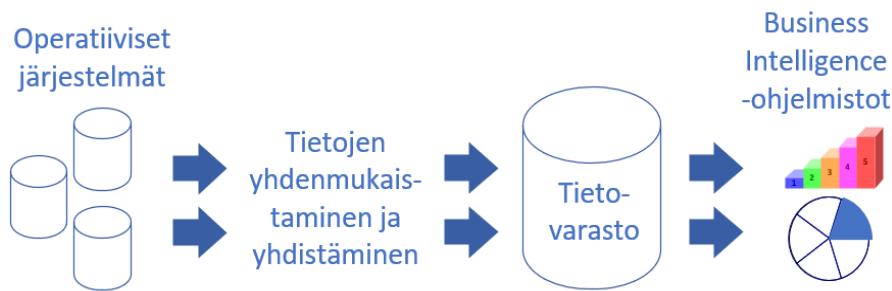
5 Yhteenveto ja johtopäätökset

5.1 Johtopäätökset

Kehitystyön tavoitteena oli tuottaa spesifit työkalut kahta osin samankaltaista tarvetta varten. Nämä työkalut eli Power BI -raportit kehitettiin, mutta niiden käyttöönotto jäi puolitiehen johtuen rajapinnan puuttumisesta haluttuun tietojärjestelmään. Kehitystyön suunnitteluvaiheen ongelmana olikin puutteellinen alkukartoitus siitä, oliko tavoitteen onnistumiselle todelliset edellytykset olemassa. Kommunikaation merkitystä ei voi liikaa korostaa tässä vaiheessa, niin opinnäytetyön tekijän, toimeksiantajan kuin organisaation ulkopuolisten asianosaisten kesken. Rajapinnan toteutumiseksi olisi tarvittu organisaation ulkopuolisen toimijan myönteistä suhtautumista kehitystyön toteuttamiseen, mutta pelkän keskusteluyhteyden avaaminen toimijaan vei kuukausia. Tämä toimija olisi tullut tunnistaa yhdeksi osapuoleksi kehitystyössä, ja pääsy tietoihin olisi tullut varmistaa sopimalla, kuten Lukka (2001) suosittelee.

Kun rajapinnan toteutuminen alkoi näyttää epätodennäköiseltä opinnäytetyön aikaikkunassa, päädyttiin kehitystyöryhmän kesken luopua tietokantayhteyden toteuttamisesta alkuperäisen idean mukaan ja kehitystyön fokus siirrettiin raporttien suunnitteluun ja tietosuojan asianmukaiseen toteutukseen. Myös koko prosessin uudelleenkäynnistämistä harkittiin, mutta näiden kahden spesifin työkalun eteen tehtyä työtä oli jo kertynyt ja kokonaan uuden aiheen valinta nähtiin riskinä opinnäytetyön toteutumiselle. Mikäli rajapinta joskus toteutuisi, työkalut olisivat kuitenkin suhteellisen helposti implementoitavissa uudelle yhteydelle. Kyselyihin ja muuttujiin kohdistuvaa teknistä työtä olisi, mutta itse kehitystyön tulos, eli tieto siitä, minkälaisista ja -muotoisista Power BI -raporteista esihenkilöt hyötyisivät, olisi valmiina. Tämän lisäksi toisen työkalun edellyttämä tietosuojan vaikutustenarviointi olisi riskianalyyseineen suurimmaksi osaksi myös jo valmiina, joskin mahdollisen uuden tietokantayhteyden toteutusta tulisi tarkastella vielä tietosuojan ja -turvan näkökulmasta.

Jo opinnäytetyösuunnitelmassa tunnistettiin tietojen hajallaan oleminen eri järjestelmissä. Kehitystyön suunnitteluvaiheessa tunnistettiin myös raportointikannan puuttuminen. Jälki- viisaasti voidaan todeta, että raportoinnin pitkän tähtäimen kehitystä ajatellen kehitystyö olisi pitänyt aloittaa eri operatiivisten järjestelmien tietojen yhteen paikkaan keräämisestä ja yhdenmukaistamisesta. Operatiiviset järjestelmät sopivat huonosti raportointiin ja analysointiin. Tietojen hyödyntäminen BI-ohjelmistoilla onnistuukin parhaiten tietokannalla, joka on tätä varten suunniteltu ja rakennettu. (Hovi ym. 2009, 14; Kavanagh & Johnson 2018, 41.) Alla kuvatussa prosessissa tiedot kertyvät eri operatiivisten järjestelmien tietokantoihin, joissa niitä myös pidetään ajan tasalla (Kuvio 6).



Kuvio 6. Tietovuo operatiivisista järjestelmistä BI-ohjelmistoihin (mukailtu Hovi ym. 2009, 14)

ETL-vaiheessa (Extract – Transform – Load) operatiivisten järjestelmien tiedot luetaan ja tallennetaan tietovarastoon. Tiedot tallennetaan yhdenmukaisessa muodossa lukemista ajatellen, ja päällekkäisiä tietoja pyritään yhdistämään ja niiden oikeellisuuteen voidaan tehdä tarkistuksia. Tietovaraston tiedot päivittyvät operatiivisista järjestelmistä tulevien muutosten myötä, mutta eivät toisinpäin. (Hovi ym. 2009, 14.) Tietovarasto (eng. Data Warehouse) on siis tietojen raportointia ja analysointia varten suunniteltu kanta, josta BI-ohjelmistoilla on helppo hakea tietoja (Tableau 2022a).

Kehitystyön viimeisen iteraation toteutukset vastaavat kukin omalla tavallaan opinnäytetyön tavoitteista johdettuihin tutkimuskysymyksiin. Esihenkilöiden palautetilaisuuden perusteella osaamisen kehittämisen raportti vastasi heidän esittämään tarpeeseensa, eli siihen, että he pystyisivät seuraamaan henkilöstön kehittämiseen tarkoitetun resurssin toteutumista toimiala- yksikkö- ja työntekijätasolla. Myös sairauspoissaolojen määrällinen seuraaminen mahdollistettiin omalla raportillaan. Sairauspoissaolojen seuraamiseen liittyneet alikysymykset, eli työhyvinvointihankkeiden kohdentamista tukevan tiedon löytäminen ja aikamääreisiin liittyvien sairauspoissaolotrendien löytäminen, eivät ole yhtä yksinkertaisia vastata. Sairauspoissaolojen raportti vastaa niihin osin. Työhyvinvointihankkeiden kohdentamiseen ei todennäköisesti riitä pelkkä määrällinen tieto sairauspoissaolojen kasvusta jossain tiettyssä kohteessa. Kohdentamisen tueksi tarvitaan paljon muutakin informaatiota, josta osa on sellaista, jota työnantajalla ei välttämättä edes ole mahdollista saada. Aikamääreisiin liittyvien sairauspoissaolotrendien löytämiseksi tehtiin kehitystyön aikana useita eri toteutuksia, mutta ainoastaan vuosivertailu päätyi viimeiseen versioon. Siten ainut aikaan liittyviä trendejä tarkasteleva raportti tarkastelee ainoastaan kahden eri vuoden eri kuukausien sairauspoissaoloja, ja se vastaa esimerkiksi kysymykseen ”kuinka paljon enemmän tai vähemmän viime vuoden keväällä sairastettiin verrattuna tähän vuoteen?”. Muita trendejä sillä ei kuitenkaan ole mahdollista löytää. Muiden trendien löytämisen mahdollistaneet visualisoinnit jäivät kehitystyön aikana pois joko siksi, että aineiston perusteella trendejä ei ollut löydettävissä, tai siksi, että ilmiötä ei pidetty merkityksellisenä työnantajan kannalta.

Tämän opinnäytetyön konstruktiot olivat aiheeltaan nimenomaisesti niitä, joille REDUn esihenkilöt olivat esittäneet sisäisessä kyselyssä toiveita ajantasaisesta Power BI -raportoinnista, ja tavoitteet olivat niihin nähden loogisia. Mikäli tiedolla johtamista tukeva kehityssuunta jatkuu REDUssa ja toimiva henkilöstöresurssin raportointikanta-tietovaranto saadaan tulevaisuudessa käyttöön, ovat nämä raportit omalta osaltaan lähes valmiita käyttöön-ottoon.

5.2 Jatkokehittäminen

Tämän opinnäytetyön puitteissa tapahtunut kehitystyö ei saavuttanut kaikkia opinnäytetyösuunnitelmassa esiteltyjä tavoitteitaan. Tuotettuja raportteja ei jaettu käyttöön REDUn esihenkilöille, koska raporteilla ei ollut opinnäytetyön kehitystyön päättämisen aikaan toimivaa tietokantayhteyttä. Nykytoteutuksellaan raporttien päivittäminen olisi vaatinut manuaalista työtä, joka sisälsi kohonneen tietosuojariskin. Jotta kehitystyö saavuttaisi viimeisen tavoitteensa, eli sen, että raportit olisivat jaettuna niiden käyttäjille, tulisi rajapintaongelma ratkaista tuotantokannan toimittajan kanssa. Tällaisen raportointikannan, eli tässä tapauksessa henkilöstöresurssin yhteisen tietovarannon kehittäminen loisi hyvät edellytykset ajantasaisen raportoinnin jatkokehitykselle REDUssa. Toimiva tietovarasto on perusedellytys jotta tietoja voidaan kunnolla hyödyntää raporteissa ja analyyseissä (Hovi ym. 2009, 14).

Raportointikannan kokoaminen henkilöstöpalveluiden eri tietojärjestelmistä mahdollistaisi monien muidenkin mittarien kehittämisen tässä opinnäytetyössä esiteltyjen lisäksi. Tällaisia voisivat olla esimerkiksi terveysperustaisten poissaolojen kustannukset, työterveyshuollon kustannukset, henkilöstön vaihtuvuus, turvallisuusilmoitukset, sisäilmailmoitukset, työtapaturmat ja työmatkatapaturmat. Myös henkilöstökyselyiden tietojen sisällyttäminen tällaiseen raportointikantaan voisi tuoda ilmi muutoksia, jotka tapahtuvat ajan kanssa. Henkilöstön sitoutuneisuuden mittaaminen voi oikein muotoiltuna tarjota vastauksia myös siihen, miksi työntekijät ovat tai eivät ole sitoutuneita työhönsä (Kavanagh & Johnson 2018, 468).

Esihenkilöiden palautetilaisuudessa annettu kommentti koskien sairauspoissaolojen raportointia toi esille oleellisen puutteen tiedoissa. Esihenkilöiden on vaikea reagoida proaktiivisesti sairauspoissaoloihin mikäli niiden syistä ei ole tietoa. Koska työnantaja ei mainittua tietoa voi kuitenkaan kerätä, tulisikin miettiä muita mahdollisia mittareita työhyvinvoinnin arviointiin. Tällaisia ovat esimerkiksi henkilöstön vaihtuvuutta ja työilmapiiriä ilmentävät mittarit (Viitala 2009, 329). Vaihtuvuutta mitattaessa mielenkiinto kohdistetaan tyypillisesti tapauksiin, joihin työnantaja voisi toiminnallaan vaikuttaa. Tällöin rajataan pois ainakin eläköitymiset ja määräaikaaisuudet, ja puhutaan niin kutsutusta nettovaihtuvuudesta. Bruttovaihtuvuus kattaa kaikki vaihtuvuustapaukset. (Mankki 2018.) Työilmapiiriä mitataan perinteisesti työtyytyväisyyskyselyillä, kuten REDUssakin tehdään. Kyselyiden tuloksia käsitellään

kyselyiden jälkeen, mutta eri vuosien kyselyiden vertailuun ei ole valmiita työkaluja. Turvalisuuksilmoitukset, sisäilmailmoitukset ja työtapaturmat voivat myös omalta osaltaan kuvastaa työhyvinvoinnin tilaa organisaatiossa, jolloin niiden ajantasainen seuranta on perusteltua.

5.3 Pohdinta

Konstruktiiiviseen tutkimusotteen ydinkäsite on uusi konstruktio, joka tämän kehitystyön yhteydessä määriteltiin tietynlaiseksi artefaktiksi jo suunnitteluvaiheessa. Hyvinkin spesifisti määritelty tavoite saattoi omalta osaltaan vaikuttaa siihen, että konstruktiolelle nähtiin vain yksi toteuma, tämän opinnäytetyön tavoitteen mukainen toteuma. Konstruktiolella on kuitenkin aina muitakin mahdollisia lopputulemia. Mikäli kehitystyö olisi aloitettu kartoittamalla edellytyksiä tavoitteen toteutumiselle, olisi kehitystyön suunnaksi saattanut muodostua jotain täysin muuta, esimerkiksi raportointikannan kehittäminen ja konstruktioleksi raportointikannan määrittely. Raportointikanta-konstruktion jälkeen jatkokehitys oltaisiin voitu suunnata raportointikannan päälle rakennettavien raporttien kehittämiseen, joista tämän opinnäytetyön artefaktit olisivat olleet yksi mahdollinen osa. Todennäköisesti raporttien suunnitteluun ei olisi päästy yhden opinnäytetyön aikana ollenkaan, mutta jatkokehittämiselle olisi ollut enemmän mahdollisia suuntia, koska raportointikanta olisi oletettavasti sisältänyt huomattavasti enemmän henkilöstöresurssin tietoja kuin mitä tarvittaisiin pelkästään sairauspoissaolojen ja osaamisen kehittämisen resurssin raportoinnin toteutuksiin. Ratkaisumalli ja ongelman ratkaiseva konstruktio olisi mahdollisesti ollut parempi pitää avoimempina aina siihen asti, kunnes opinnäytetyöntekijä olisi kerennyt saavuttamaan syvällisen yleisnäemyksen kohdeorganisaation lähtötilanteesta ja ongelma-alue olisi voitu käsitteellistää siten, että osapuolten välinen viestintä olisi toiminut paremmin.

Kun kehitystyöryhmän kanssa lopulta päästiin kehittämään varsinaista käytännön toteutusta, innovointi ja erilaisten ratkaisujen kokeilu toimi sulavasti. Lukka (2001) kuvaa projektin tätä vaihetta luonteeltaan luovaksi ja heuristiseksi, ja raporttien eri iteraatiot erosivatkin toisistaan melko paljon. Kehitystyöryhmän sisäinen viestintä toimi ideoinnissa hyvin ja ketterästi, ja toimiva kommunikaatio olikin avainasemassa konstruktion etsiessä lopullista muotoaan. Tekniset haasteet tarjosivat paljon edellytyksiä kokemukselliselle oppimiselle.

5.4 Kehittämistyön eettisyys ja luotettavuus

Opinnäytetyö on tehty Ammattikorkeakoulujen rehtorineuvosto Arene Ry:n opinnäytetöitä koskevia eettisiä ohjeita noudattaen (Arene 2020). Kehitystyön eri vaiheissa käsiteltiin henkilötietoja, ja näiden henkilötietojen suojaamisessa noudatettiin REDUn henkilötietojen käsittelyn periaatteita. Näihin periaatteisiin kuuluu muun muassa esimerkiksi toimiminen

tietosuoja-asioissa. Tietosuojalainsäädännön noudattaminen, tietosuojan varmistaminen tietoturvaisilla tietojärjestelmillä ja tietosuojan jatkuva parantaminen olivat siten mukana koko prosessin ajan.

Opinnäytetyön tietosuojan voi nähdä koskeneen kahta eri asiaa: opinnäytetyön kehitysosassa käytetyssä aineistossa olleita todellisia henkilötietoja, sekä kehitysosan kehittämän uuden käsittelytoiminnon tulevaisuudessa käsittelemiä henkilötietoja, joita varten tietosuojan vaikutustenarviointi ja riskianalyysi tehtiin. Kehitystyön alkuvaiheessa käsiteltujen todellisten henkilötietojen käsittely pyrki vain yhteen asiaan eli henkilötietojen anonymisointiin. Henkilötiedot anonymisoitiin, ja käsittelytoimintojen jatkokehitys tehtiin näillä anonymisoiduilla tiedoilla. Anonymisoituja tietoja ei enää katsota henkilötiedoiksi, eikä niihin näin ollen sovelleta tietosuojasäännöksiä. (Tietosuojavaaltuutetun toimisto 2022g.)

Huomioitavaa on myös, että tutkimus ei kohdistunut varsinaiseen aineistoon. Aineiston henkilöt eivät olleet tutkimuksen kohteena, eivätkä myöskään aineiston muut määrälliset tai laadulliset ominaisuudet. Aineistoa käytettiin kehittämään uutta käsittelytoimintoa, ja käsittelytoimintojen kehittäminen oli tutkimuskohde. Kehittämisessä käytettiin anonymisoitua oikeaa tietoa, jotta kehitystyössä käytetty data olisi muodoltaan mahdollisimman lähellä tietolähteistä saatavaa oikeaa dataa. Aineistoa hankittaessa aineiston käytölle haettiin REDUn tutkimuslupa. Tutkimusluvassa määriteltiin käsiteltävät henkilötiedot ja salassa pidettävät tiedot. Aineisto tuhottiin kehitystyön päätyttyä.

Kehittämistyön luotettavuutta arvioidaan kiinnittämällä huomiota käytettyjen menetelmien oikeellisuuteen ja tarkoituksenmukaisuuteen. Kehittämistyön tuloksena syntyvät muutokset ovat myös arvioinnin kohteita. Arviointi perustuu kehitystyön dokumentaatioon. (Hyväri & Vuokila-Oikkonen 2020.) Tutkimuksellisella kehittämistyöllä voidaan pyrkiä ratkaisemaan käytännöstä nousseita ongelmia tai uudistamaan työelämän käytäntöjä. Opinnäytetyössä pyrittiin luomaan uusia työkaluja esihenkilöille, ja työkalujen oli tarkoitus ratkaista käytännön ongelmia kuten tiedon saavutettavuuteen liittyviä haasteita. Mikäli kehitystyö olisi päässyt täyteen tavoitteeseensa, ja työkalut olisi otettu REDUssa käyttöön, olisivat ne uudistaneet omalta osaltaan REDUn käytäntöjä ja käytettävyyttä tiettyjen henkilöstöresurssin tietojen suhteen. Siltä osin tutkimuksellista kehittämistyötä voidaan pitää onnistuneena ja järkevänä valintana käytetyksi menetelmäksi. Niin ikään konstruktiivisen tutkimusotteen valintaa voidaan pitää perusteltuna, koska opinnäytetyösuunnitelmassa tavoitteeksi määriteltiin tietynlainen konstruktio. Konstruktio käsitteenä sisältää myös toteuttamisyrityksen, jolla testataan sen käytännön soveltuvuutta (Lukka 2001). Käytännön soveltuvuuteen hankittiin näkemyksiä kehitystyön lopputuotteen oletetuilta käyttäjiltä eli REDUn esihenkilöiltä. Vaikka palaute oli konstruktion käyttöönottoon kannustavaa, olisi validoinnin tueksi ollut hyvä hankkia

dokumentaatiota esimerkiksi palautelomakkeen muodossa. Muulta osin opinnäytetyöntekijä on pyrkinyt dokumentoimaan kehitystyöhön vaikuttaneet asiat kirjallisesti siltä osin aina kuin se on ollut mahdollista tutkimusluvan ja tietosuojan puitteissa.

5.5 Oman oppimisen ja ammatillisen kehittymisen tarkastelut

Opinnäytetyön tekijän henkilökohtaisena oppimistavoitteena oli oman ict-ammattiosaamisen syventämisen lisäksi hankkia uusia tietoja ja taitoja data-analytiikasta, raportoinnista ja aineettoman pääoman mittaamisesta. Näiden lisäksi opinnäytetyö tarjosi oppimismahdollisuuksia projektinhallintaan ja suunnitteluun. Prosessi ei edennyt virheittä, mutta näiden virheiden objektiivinen tarkastelu mahdollisti oppimista etenkin kehittämistyön suunnittelun suhteen, mitä voidaan pitää hyödyllisenä taitona asiantuntijatyössä.

Opinnäytetyöntekijän opinnäytetyötä edeltänyt osaamistaso analytiikan, raportoinnin ja tietosuojan suhteen oli matala. Opinnäytetyön myötä tiedot ovat lisääntyneet huomattavasti, aina pienistä yksityiskohtaisista asioista kuten kyselykielten osaamisesta isoihin kokonaisuuksiin kuten tietosuojaan. Tietosuojan osuus ja merkitys opinnäytetyössä olikin kuviteltua suurempi. Vaikutustenarvioinnin ja tasapainotestin konkreettinen tekeminen pilkkoivat isoa ja monimutkaista aihetta helpommin sisäistettäviin kokonaisuuksiin. Kehitystyön eri iteraatioiden aikainen PowerBI -raporttien ideointi, toteuttaminen ja viimeistely lisäsivät näkemystä tietojen oikeellisuuden varmistamisen tärkeydestä, aina tietolähteestä viimeiseen raporttiin asti.

Teoreettiseen viitekehikseen perehtyminen toi lisäymmärrystä etenkin aineettoman pääoman ja henkilöstöresurssin käsitteisiin, jota ilman niistä kertovaa raportointia olisi hankala rakentaa. Konstruktiviseen tutkimusotteeseen kiinteästi liittyvää kokemuksellista oppimista tapahtui etenkin eri iteraatioita tuottaessa, mutta myös eri sidosryhmien välisessä viestinnässä - etenkin silloin kun viestintä oli haastavaa.

Ylempään ammattikorkeakoulututkintoon johtavan koulutuksen myötä opinnäytetyöntekijällä on herännyt kiinnostus hyödyntää opittuja taitoja vaativammissa ja monimutkaisemmissa työtehtävissä. Osin tämä onkin jo tuottanut hedelmää uusien asiantuntijatehtävien myötä. Uuden osaamisen hankkiminen on nostanut opinnäytetyöntekijän mielenkiintoa myös tutkimusta kohtaan.

Lähteet

- Agendum. 2019. Tietosuojaselosteen vinkkilista. Viitattu 27. 3.2022. Saatavissa <https://akatemia.digiturvamalli.fi/tietosuojaseloste>
- Arene. 2020. Ammattikorkeakoulujen opinnäytetöiden eettiset suositukset. Viitattu 27.3.2022 Saatavissa <https://www.arene.fi/wp-content/uploads/Raportit/2020/AMMATTI-KORKEAKOULUJEN%20OPINN%C3%84YTET%C3%96IDEN%20EETTISET%20SUOSITUKSET%202020.pdf?t=1578480382>
- Armstrong, M. 2006. A Handbook of Management Techniques. 3. painos. Kogan Page.
- Baesens, B. 2014. Analytics in a Big Data World. New Jersey: John Wiley & Sons, Inc.
- Banfield, P.; & Kay, R. 2012. Introduction to Human Resource Management. New York: Oxford University Press Inc.
- Bishop, T. 2021. 'Unprecedented' Microsoft Azure database vulnerability impacts thousands of customers, researchers warn. Viitattu 27.3.2022 Saatavissa <https://www.geek-wire.com/2021/unprecedented-microsoft-azure-database-vulnerability-impacts-thousands-customers-researchers-warn/>
- Emery, J. 2020. Five Ways to Mislead with Data Visualizations. Viitattu 15.11.2021 Saatavissa <https://www.tessellationtech.io/top-five-ways-mislead-data-visualization/>
- Euroopan komissio. 2022. Mitkä tiedot ovat henkilötietoja? Viitattu 27.3.2022 Saatavissa https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_fi
- Heikkinen, S. 2019. Tilastojen seitsemän kuolemansyntiä: Näin sinua harhautetaan. YLE. Viitattu 20.11.2021 Saatavissa <https://yle.fi/aihe/artikkeli/2019/04/09/tilastojen-seitsemän-kuolemansyntia-nain-sinua-harhautetaan>
- Hovi, A., Hervonen, H., & Koistinen, H. 2009. Tietovarastot ja Business Intelligence. Porvoo: WSOY.
- Humanistinen ammattikorkeakoulu. 2021. Opinnäytetyöopas YAMK. Viitattu 26.11.2022 Saatavissa <https://humak.libguides.com/c.php?g=688355&p=4925415>
- Hyväri, S. & Vuokila-Oikkonen, P. 2020. Osallistavan ja tutkivan kehittämisen opas 2.0. Viitattu 8.4.2022. Saatavissa <https://libguides.diak.fi/c.php?g=670543&p=4760642#kehi>
- IBM Cloud Education. 2020. Data Modeling. Viitattu 1. 2.2022. Saatavissa <https://www.ibm.com/cloud/learn/data-modeling#toc-types-of-d-Fne2rmPT>
- Innowise. 2019. Evästeet nettisivuilla ja GDPR - miten pitää ilmoittaa? Viitattu 27.3.2022 Saatavissa <https://www.innowise.fi/fi/evasteet-eli-cookies-nettisivuilla-ja-gdpr/>
- Jalonen, H. 2015. Tiedolla johtamisen näyttämö ja kulissit. Vaasan yliopisto. Saatavissa https://www.researchgate.net/profile/Harri_Jalonen/publication/275329037_Tiedolla_johtamisen_nayttamo_ja_kulissit/links/5bf18d7c4585150b2bc01d48/Tiedolla-johtamisen-naeyt-taemoe-ja-kulissit.pdf
- JAMK. 2021. Opinnäytetyön ohjaajan käsikirja. Viitattu 26.3.2022. Saatavissa <https://oppi-materiaalit.jamk.fi/yamk-kasikirja/tyoelaman-tutkiva-kehittamistoiminta/>
- Johns Hopkins University of Medicine. 2021. Coronavirus Resource Center: Finland Overview. Viitattu 1.11.2021. Saatavissa <https://coronavirus.jhu.edu/region/finland>

- Julkaisut Helsinki. 2020. Työhyvinvointi: Työhyvinvointia johdetaan tiedon avulla. Viitattu 26.3.2022 Saatavissa <https://julkaisut.hel.fi/fi/julkaisut/henkilostoraportti-2020/tyohyvinvointi-tyohyvinvointia-johdetaan-tiedon-avulla>
- Kaur, R. 2021. Power BI vs Tableau vs Qlik Sense: Which BI Tool is the Winner? Viitattu 26.3.2022 Saatavissa <https://www.selecthub.com/business-intelligence/tableau-vs-qlikview-vs-microsoft-power-bi/>
- Kavanagh, M. J. & Johnson, R. D. 2018. Human resource information systems: basics, applications, and future directions. Los Angeles: SAGE.
- Kesti, M. 2010. Strateginen henkilöstötuottavuuden johtaminen. Helsinki: Talentum.
- Kosonen, M. 2019. Tiedolla johtamisen käsikirja. Xamk Kehittää 81. Kaakkois-Suomen ammattikorkeakoulu. Viitattu 26.3.2020. Saatavissa <https://digitalia.xamk.fi/tijo>
- KT Kuntatyönantajat. 2013. Henkilöstövoimavarojen arviointi. Suositus henkilöstöraportoinnin kehittämiseen. Viitattu 8.4.2022. Saatavissa <https://www.kt.fi/sites/default/files/media/document/a-henkilostovoimavarojen-arviointi-2013.pdf>
- Kujala, T. 2005. Tietokantojen perusteet (58132), KEVÄT 2005. Helsingin yliopisto. Viitattu 5.12.2021 Saatavissa <https://www.cs.helsinki.fi/u/tkuujala/tikapeK2005/pdf/relaatiomalli.pdf>
- Kujansivu, P., Lönnqvist, A., Jääskeläinen, A. & Sillanpää, V. 2007. Liiketoiminnan aineetomat menestystekijät. Helsinki: Talentum Media Oy.
- Kunta- ja hyvinvointialueyönantajat KT. 2015. Kunnallisen henkilöstön osaamisen kehittämistä koskeva suositus sekä työ- ja virkaehtosopimus ammattiyhdistyskoulutuksesta. Viitattu 12.12.2021. Saatavissa <https://www.kt.fi/henkilostojohtaminen/suosituks/osaamisen-kehittaminen>
- Kunta- ja hyvinvointialueyönantajat KT. 2017. Osaamisen johtaminen on osa strategista henkilöstöjohtamista. Viitattu 26.11.2021 Saatavissa <https://www.kt.fi/henkilostojohtaminen/osaamisen-johtaminen>
- Laihonen, H. & Lönnqvist, A. 2013. Tiedolla johtaminen tarkoittaa tiedon hyödyntämistä. Tietoasiantuntija. Saatavissa <https://tietovirta.wordpress.com/2013/11/06/tiedolla-johtaminen-tarkoittaa-tiedon-hyodyntamista/>
- Laihonen, H., Hannula, M., Helander, N., Ilvonen, I., Jussila, J., Kukko, M., Kärkkäinen, H., Lönnqvist, A., Myllärniemi, J., Pekkola, S., Virtanen, P., Vuori, V., Yliniemi, TT. 2013. Tietojohtaminen. Tampere: Tampereen teknillinen yliopisto.
- Lukka, K. 2000. The key issues of applying the constructive approach to field research. Turun yliopisto. Viitattu 10.10.2021. Saatavissa https://www.researchgate.net/publication/281549256_The_key_issues_of_applying_the_constructive_approach_to_field_research
- Lukka, K. 2001. Konstruktiivinen tutkimusote. Viitattu 27.11.2021. Saatavissa <https://metodix.fi/2014/05/19/lukka-konstruktiivinen-tutkimusote/>
- Mankki, M. 2018. Viikon kysymys: Miten henkilöstön vaihtuvuutta mitataan? Elinkeinoelämän keskusliitto. Viitattu 27. Maaliskuu 2022. Saatavissa <https://ek.fi/ajankohtaista/blogit/viikon-kysymys-miten-henkiloston-vaihtuvuutta-mitataan/>

- Microsoft. 2022a. Microsoft Docs. Mikä Power BI on? Viitattu 26.3.2022. Saatavissa <https://docs.microsoft.com/fi-fi/power-bi/fundamentals/power-bi-overview>
- Microsoft. 2022b. What is business intelligence? Viitattu 26.3.2022. Saatavissa <https://powerbi.microsoft.com/en-us/what-is-business-intelligence/>
- Minilex. 2021. Yksityisyyden suoja Suomessa. Viitattu 27.3.2022. Saatavissa <https://www.minilex.fi/a/yksityisyyden-suoja-suomessa>
- Minilex. 2022. Mikä on henkilörekisteri? Viitattu 27. 3. 2022. Saatavissa <https://www.minilex.fi/a/mik%C3%A4-on-henkil%C3%B6rekisteri>
- Ojasalo, K., Moilanen, T., & Ritalahti, J. 2015. Kehittämistyön menetelmät. Uudenlaista osaamista liiketoimintaan. Helsinki: Sanoma Pro Oy.
- OpiTietosuoja. 2020. Yleistä tietosuojasta. Viitattu 1.2.2022. Saatavissa <https://opitietosuoja.fi/fi/aloitus/tietosuoja>
- Power, D. J. 2014. Using 'Big Data' for analytics and decision support. Teoksessa Journal of Decision Systems 23, 222-228.
- Power, D. J. 2016. Data science: supporting decision-making. Journal of Decision Systems 25, 345-356.
- REDU. 2021. Tietosuoja REDU-konsernissa. Saatavissa <https://www.redu.fi/fi/REDU/Tietosuoja>
- REDU. 2022a. Konserni ja organisaatio. Viitattu 27.3.2022 Saatavissa <https://www.redu.fi/fi/REDU/konserni/organisaatio>
- REDU. 2022b. Strategia. Viitattu 27.3.2022 Saatavissa <https://www.redu.fi/fi/redu/konserni/strategia>
- Roberts, B. 2012. Data Data Everywhere and Not a Drop of Value. Bryce Dot VC blogi. Saatavissa <https://bryce.vc/post/15300645787/data-data-everywhere-and-not-a-drop-of-value>
- Saaranen-Kauppinen, A. & Puusniekka, A. 2022. Tapaustutkimus. Viitattu 8.4.2022. Saatavissa https://www.fsd.tuni.fi/menetelmaopetus/kvali/L5_5.html
- Suominen, S. & Suominen, J. 2015. Laatu raportointiin Excelillä. Saarijärvi: Saarijärven Offset Oy.
- Tableau. 2022a. Understanding the Value of Data Warehousing & Business Intelligence (BI). Viitattu 2.4.2022. Saatavissa <https://www.tableau.com/learn/articles/value-of-bi-data-warehousing>
- Tableau. 2022b. What Is Data Visualization? Definition, Examples, And Learning Resources. Viitattu 26. Maaliskuu 2022 Saatavissa Tableau: <https://www.tableau.com/learn/articles/data-visualization>
- Tieteen termipankki. 2016. Soveltava tutkimus. Viitattu 26.3.2022 Saatavissa [https://tieteen termipankki.fi/wiki/Filosofia:soveltava tutkimus](https://tieteen termipankki.fi/wiki/Filosofia:soveltava_tutkimus)
- Tietojohtaminen Ry. 2015. @KMspecialists. Twiitti. Viitattu 26.3.2022. Saatavissa <https://twitter.com/KMspecialists/status/595211798887976962>
- Tietosuojalaki 5.12.2018/1050.

Tietosuojavaltuutetun toimisto. 2020. Työelämän tietosuojan käsikirja. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/documents/6927448/8214540/Ty%C3%B6el%C3%A4m%C3%A4n+tietosuoja+k%C3%A4sikirja+2020-Tietosuojavaltuutetun+toimisto.pdf>

Tietosuojavaltuutetun toimisto. 2021. Tietosuoja. Viitattu 8.4.2022. Saatavissa <https://tietosuoja.fi/tietosuoja>

Tietosuojavaltuutetun toimisto. 2022a. Automaattinen päätöksenteko ja profilointi. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/automaattinen-paatoksenteko-profilointi>

Tietosuojavaltuutetun toimisto. 2022b. Erityisten henkilötietoryhmien käsittely. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/erityisten-henkilotietoryhmien-kasittely>

Tietosuojavaltuutetun toimisto. 2022c. Henkilötietojen käsittelijän velvollisuudet. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/henkilotietojen-kasittelijan-velvollisuudet>

Tietosuojavaltuutetun toimisto. 2022d. Mikä on henkilötieto? Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/mika-on-henkilotieto>

Tietosuojavaltuutetun toimisto. 2022e. Milloin henkilötietoja saa käsitellä? Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/kasittelyperusteet>

Tietosuojavaltuutetun toimisto. 2022f. Osoita noudattavasi tietosuojasäännöksiä. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/osoitusvelvollisuus>

Tietosuojavaltuutetun toimisto. 2022g. Pseudonymisoidut ja anonymisoidut tiedot. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/pseudonymisointi-anonymisointi>

Tietosuojavaltuutetun toimisto. 2022h. Rekisterinpitäjän oikeutettu etu. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/rekisterinpitajan-oikeutettu-etu>

Tietosuojavaltuutetun toimisto. 2022i. Rekisteröidyn suostumus. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/rekisteroidyn-suostumus>

Tietosuojavaltuutetun toimisto. 2022k. Vaikutustenarviointi. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/vaikutustenarviointi>

Tietosuojavaltuutetun toimisto. 2022l. Henkilötietojen käsittely. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/henkilotietojen-kasittely>

Tietosuojavaltuutetun toimisto. 2022m. Seloste käsittelytoimista. Viitattu 27.3.2022. Saatavissa <https://tietosuoja.fi/seloste-kasittelytoimista>

Titus. 2021. Data Privacy vs. Data Security: What's the Difference? Blogi. Viitattu 27.3.2022. Saatavissa <https://www.titus.com/blog/data-security/data-privacy-vs-data-security-whats-the-difference>

Työsuojeluhallinnon verkkopalvelu. 2021. Terveystietotiedot. Viitattu 27.3.2022. Saatavissa <https://www.tyosuojelu.fi/tyosuhde/oikeudet-ja-velvollisuudet-tyossa/yksityisyyden-suoja/terveystiedot>

Verohallinto. 2021. Voit vastustaa vuoden 2020 tuloverotietojen sähköistä luovutusta tiedotusvälineille. Viitattu 27.3.2022. Saatavissa <https://www.vero.fi/tietoa-verohallinnosta/verohallinnon-esittely/tietosuoja-ja-julkisuus/verotuksen-julkiset-tiedo/henkiloasiakkaan-tuloverotuksen-julkise/voit-vastustaa-vuoden-2020-tuloverotietojen-s%C3%A4hk%C3%B6ist%C3%A4-luovutusta-tiedotusv%C3%A4line>

Viitala, R. 2009. Henkilöstöjohtaminen - Strateginen kilpailutekijä. Helsinki: Edita Publishing Oy.

Wikipedia. 2020. Tietojohtaminen. Viitattu 26.3.2022. Saatavissa <https://fi.wikipedia.org/wiki/Tietojohtaminen>

Wikipedia. 2021. Aineeton pääoma. Viitattu 26.3.2022. Saatavissa https://fi.wikipedia.org/wiki/Aineeton_p%C3%A4%C3%A4oma

Wikipedia. 2022. Microsoft Power BI. Viitattu 26.3.2022. Saatavissa https://en.wikipedia.org/wiki/Microsoft_Power_BI

YLE. 2021. Pandemiasta huolimatta Kelan korvaamien pitkien sairauspoissaolojen määrä väheni viime vuonna. Viitattu 26.3.2022. Saatavissa <https://yle.fi/uutiset/3-11970314>

Zheng, J. 2022. Data Visualization for Analytics and Business Intelligence - A Comprehensive Overview. Viitattu 26.3.2022. Saatavissa https://www.researchgate.net/publication/327578825_Data_Visualization_for_Analytics_and_Business_Intelligence_A_Comprehensive_Overview

Österberg, M. 2005. Henkilöstöasiantuntijan käsikirja. Helsinki: Edita Publishing Oy