



Datasta tilannekuvaksi

TyösyRJintään liittyvät asiakasyhteYdenotot
työsuojeluviranomaiselle 2023

Insinööri (YAMK) opinnäytetyö
Tietojohtaminen ja älykkäät ratkaisut

Kevät 2024

Mira Nousiainen

Tietojohtaminen ja älykkäät palvelut

Tekijä Mira Nousiainen

Työn nimi Datasta tilannekuvaksi – Työsyrijintään liittyvät asiakasyhteydenotot työsuojeluviranomaiselle 2023

Ohjaaja Essi Isohanni

Tiivistelmä

Vuosi 2024

Tämä opinnäytetyö käsittelee työsyrijintään, tietojohtamisen ja datan laadun teemoja. Työn aikana kehitettiin raportointityökalu, jonka avulla toimeksiantaja pystyy luomaan ajantasaisen tilannekuvan työsyrijintään liittyvistä asiakasyhteydenotoista nopeasti ja vähin toimenpitein. Raportointityökalu kehitettiin soveltamalla CRISP-DM-tiedonlouhintaprosessia. Toimeksiantajan käyttöönottamassa raportointityökalussa hyödynnetään toimeksiantajan omistamaa rakenteellista dataa laajasti.

Opinnäytetyön tavoitteena oli myös luoda asiakasyhteydenoton kuvaustekstiin lisää rakennetta, jotta toimeksiantaja pystyisi tuottamaan parempaa tietoa työsyrijintään eri syrijintämuodoista. Asiakasyhteydenoton kuvausteksti oli lähtötilanteessa rakenteetonta tekstidataa. Työssä toteutettiin teorialähtöinen temaattinen sisällönanalyysi työsyrijintään liittyviin asiakasyhteydenottoihin. Sisällönanalyysin perusteella luonnosteltiin syrijintäperusteita kuvaavia hakusanoja, joiden avulla oli tarkoitus luoda syrijintäperusteiden esiintymisraportti. Hakusanojen testausvaiheessa havaittiin runsaasti vääriä positiivisia havaintoja, minkä takia syrijintäperusteiden raportoinnista luovuttiin.

Opinnäytetyön johtopäätös on, ettei asiakasyhteydenoton kuvausteksti sovellu työsyrijintään perusteiden esiintuontiin. Toimeksiantajalle esitettiin kehitysehdotuksena muun muassa yhdenmukaisen tapausluokituksen käyttöönottoa ja kirjauskäytäntöjen vakioimista. Näiden muutosten avulla toimeksiantaja pystyisi laajentamaan työsyrijintään raportoinnin koskemaan myös eri syrijintäperusteita.

Toteutettu sisällönanalyysi tuotti kokonaiskuvan vuoden 2023 työsyrijintään eri syrijintämuodoista. Kolme tyypillisintä syrijintäperustetta ovat terveydentila, muu henkilöön liittyvä syy ja alkuperä.

Opinnäytetyön toimeksiantaja on Etelä-Suomen aluehallintovirasto työsuojelun vastuualue. Tutkimusaineisto sisältää kuitenkin kaikki työsuojeluhallinnon vastaanottamat työsyrijintään liittyvät asiakasyhteydenotot vuodelta 2023. Syntynyt kokonaiskuva on siten valtakunnallinen.

Avainsanat työsyrijintä, data, datan laatu

Sivut 37 sivua ja liitteitä 3 sivua

Transformative Leadership and Smart Services

Author Mira Nousiainen

Year 2024

Subject From Data to Situation Awareness – Discrimination -Related Customer
Contacts for OSH 2023

Supervisors Essi Isohanni

This thesis is focused on discrimination at workplace, information management, and data quality. During this thesis reporting template was created for a client to form an overview of workplace discrimination customer contacts. The finished reporting template extensively utilize structured data owned by the client.

The second goal of this thesis was to improve the structure of description texts of the customer contact to derive statistical insights on discrimination grounds. The description text consists of unstructured textual data. A theory-driven thematic content analysis was conducted on customer contacts. This analysis served as the foundation for identifying keywords that describe different grounds of discrimination. The intention was to create a comprehensive report detailing the occurrence of these grounds. However, during the rigorous testing phase of these keywords, a significant number of false positive observations emerged, leading to the decision to discontinue reporting on discrimination grounds.

The conclusion of this thesis is that the existing description text in a customer contact is not suitable for reporting the grounds of discrimination at workplace. The development proposal for the client is to adapt a consistent classification system and standardization of information and data entries. This would allow the client to expand workplace discrimination reporting to cover various discrimination criteria as well.

The content analysis provided an overview of grounds of discrimination at workplace in the year 2023. The three most typical forms of discrimination identified were related to state of health, other personal reasons, and the origin of the person.

The client for this thesis was the Occupational Safety and Health Division at the Regional State Administrative Agency for Southern Finland. However, the research material was nationwide, encompassing all client contacts related to discrimination at workplace received by the Occupational Safety and Health Administration.

Keywords discrimination at workplace, data, quality of data

Pages 37 pages and appendices 3 pages

Sisällys

1	Johdanto	1
2	Kehittämistyön taustat.....	3
2.1	Työsyryjä	3
2.2	Tietojohtaminen	6
2.3	Raportointi ja analytiikka	8
2.4	Tilannekuva	9
2.5	Datan laatu automatisoinnin edellytyksenä	10
3	Tarkoitus ja tavoitteet.....	12
4	Tutkimusmenetelmät.....	13
4.1	Tiedonlouhintaprosessi CRISP- DM.....	14
4.2	Teoriapohjainen temaattinen analyysi	16
4.3	Puolistrukturoitu ryhmähaastattelu	17
5	Suunnittelu ja toteutus.....	17
5.1	Raportointityökalun määrittely.....	17
5.2	Asiakasyhteydenottojen sisällönanalyysin suunnittelu.....	19
5.3	Raportointityökalun toteutus.....	20
5.3.1	Datan laadun arviointi.....	20
5.3.2	Visuaalinen tilannekuva	22
5.3.3	Asiakasyhteydenottojen sisällönanalyysi	23
5.3.4	Rakenteisuuden lisääminen asiakasyhteydenottoja koskevaan dataan	26
5.4	Kehitystyö tulosten yhteenveto.....	27
5.5	Kehitystyö toteutuksen arviointi.....	30
6	Pohdintaa ja kehitysehdotuksia	31
7	Johtopäätökset.....	35
	Lähteet	38

Kuvat, taulukot ja kaavat

Kuva 1. Tietojohtamisessa päätöksenteko ohjaa tiedon keräämistä ja analyysiä (Leskelä ym., 2019, s. 50).....	7
--	---

Kuva 2. Tiedon jalostuminen siirtymissä. (Cong & Pandya, 2003, s. 26; Finto Suomalainen asiasanasto- ja ontologiapalvelu, n.d.b; Listenmaa, 2023, ss. 35–38).....	7
Kuva 3. Analytiikan tasot Sivarajah ym. (2017, s. 226) mukaillen.....	9
Kuva 4. Datan laatutekijät mukaillen Sivula ym. (2023, s. 60) ja standardi SFS-ISO/IEC 25012:2020.....	11
Kuva 5. CRISP-DM-prosessin vaiheet laajennetuin tiedon sekä ensimmäisen vaiheen riippuvuussuhteet muihin vaiheisiin (Jackson, 2002, s. 283; Sharma & Osei-Bryson, 2009, s. 4115).	14
Kuva 6. Yhteydenoton kuvaustekstin merkkimäärän vaihtelu.....	22
Taulukko 1. Kuvauksia häirinnästä (Jauhola ym., 2022, ss. 62–172)	4
Taulukko 2. MVP-tuotteen ominaisuuksia ryhmähaastattelun perusteella.	18
Taulukko 3. Esimerkki tutkimuksessa hyödynnettävästä ristiintaulukoinnista.	19
Taulukko 4. Kehitystyössä tarvittavan datan laadun arviointi.....	21
Taulukko 5. Yhteydenotosta ilmenevien syrjäntäperusteiden määrä.	24
Taulukko 6. Asiakasyhteydenotoista ilmenevien syrjäntäperusteiden tapausmäärä	25
Taulukko 7. MVP-tuotteen määritysten toteutumisen arviointi.	29

Liitteet

Liite 1.	Aineistonhallintasuunnitelma
----------	------------------------------

1 Johdanto

Suomalaisen työelämän tulisi olla reilua, yhdenvertaista, syrjimätöntä ja kaikin puolin turvallista. Näin ei kuitenkaan ole ja vuosittain lukuisat henkilöt ovat yhteydessä työsuojeluviranomaiseen selvittääkseen onko heitä syrjitty työelämässä. Tällaisia henkilöön liittyviä kiellettyjä syrjintäperusteita ovat muun muassa alkuperä, uskonto, perhesuhteet, seksuaalinen suuntautuminen, terveydentila tai ikä. Yhteiskunnallisesti on huolestuttavaa, että kaikki kielletyt syrjintäperusteet esiintyvät suomalaisessa yhteiskunnassa. (Työsuojeluhallinto, 2023)

Tässä opinnäytetyössä ei tutkita työsyrynnän esiintymisen laajuutta eikä esiintymismuotoja vaan sen sijaan työssä syvennytään työsyrynnään liittyvän raportoinnin kehittämiseen. Työssä parannetaan toimeksiantajan kyvykkyyttä työsyrynnän tilannekuvan ylläpitoon.

Työsuojeluviranomainen valvoo yhdessä yhdenvertaisuusvaltuutetun kanssa sitä, ettei syrjintää esiintyisi suomalaisessa yhteiskunnassa ja työpaikoilla. Tilanteet ja henkilön kokemukset syrjinnästä ovat hyvin subjektiivisia. Työsuojeluviranomainen selvittää asiakkaan aloitteesta sitä, onko henkilön kokemus erilaisesta kohtelusta ollut lainsäädännössä tarkoitettua kiellettyä syrjintää. Vuonna 2022 työsuojeluviranomainen kirjasi noin 540 asiakasyhteydenottoa työsyrynnästä. (Työsuojeluhallinto, 2023)

Työsuojeluhallinto laatii vuosiraportin yhdenvertaisuuden ja syrjinnän kiellon valvonnasta, jossa käsitellään työsuojeluviranomaiselle tehtyjä asiakasyhteydenottoja sekä tarkastushavaintoja. Tähän asti raportti on tehty vain kerran vuodessa vakiintuneen vuosiraportointisyklin mukaisesti. Vuosiraportointi mahdollistaa vain pitkän aikavälin muutosten havainnoinnin. Tämän opinnäytetyön ennakkoselvitysvaiheessa tunnistettiin raportointisykliin vaikuttavaksi tekijäksi se, että raportointityö toteutetaan manuaalisesti ja työ on hyvin aikaa vievää. Raportoinnin manuaalisella toteutuksella on siten haittavaikutuksia. Muun muassa tietopyyntöjen vastaukset eivät välttämättä sisällä kaikista ajantasaisinta tietoa eikä raportointisykli tue riittävästi toimeksiantajan oman toiminnan suunnittelua raportointiajankohtien välillä.

Työsuojeluviranomainen toteuttaa lakisääteistä tehtävää, joten datan analysointitarve ei liity uusien liiketoimintamallien tunnistamiseen. Työsuojeluhallinnon tarkoitus on kuitenkin osallistua yhteiskunnalliseen keskusteluun ja jakaa ajantasaista tietoa myös työsyrynnän muodoista. Tällöin nykyinen raportointisykli ei tuota tarpeeksi nopeasti tilannetietoa työelämäkentän muutoksista. Samalla jää myös hyödyntämättä datasta saatavat hyödyt

organisaation omien resurssien ohjaukseen ja toiminnan priorisointiin. (Väre, 2019, ss.46–51)

Tietojohtaminen jakautuu kattokäsitteenä kahteen kokonaisuuteen – tiedon johtamiseen sekä tiedolla johtamiseen. Tietojohtamisen tekniset näkökulmat datan hankinnasta, varastoinnista, siirtämisestä ja suojaamisesta sekä kirjaamiskäytännöistä tietojärjestelmiin muodostavat tiedon johtamisen kokonaisuuden. Tiedon johtaminen keskittyy siis tietovirtojen hallintaan. Tiedon johtaminen kehittää tiedon hyödyntämisen edellytyksiä ohjaamalla tietovirtoja ja huolehtimalla tallennettavan datan laadusta. Organisaation datasta jalostetun informaation perusteella tehtyjen johtopäätösten esittäminen sekä tiedon hyödyntämisen mahdollistaminen päätöksenteossa muodostaa tiedolla johtamisen kokonaisuuden. (Leskelä ym., 2019, s. 24, 2019, s. 16; Listenmaa, 2023, ss. 45–45)

Tässä opinnäytetyössä keskitytään toimeksiantajan hallussa olevan datan jalostamiseen informaatioksi ja tietämykseksi. Tavoitteena on tuottaa toimeksiantajalle helposti käytettävä raportointityökalu, jolla työsyrijntää koskevien asiakasyhteydenottojen sisältöä koskeva visuaalinen tilannekuva saadaan muodostettua ja päivitettyä. Näitä asioita lähestytään tutkimuskysymyksillä:

- Miten toimeksiantajan raportointitehokkuutta voidaan parantaa?
- Kuinka työsyrijntän kokonaistilanne saadaan selvitettyä?

Tutkimuskysymyksiä tarkastellaan toimeksiantajan hallussa olevan datan kautta.

Työsyrijntään liittyy myös kestävän kehityksen näkökulmia. YK:n Agenda 2030-toimintaohjelma sisältää tavoitteita, joilla pyritään turvaamaan kestävä kehitys yhteiskunnan eri osa-alueilla. Agendan viides tavoite ”saavuttaa sukupuolten välinen tasa-arvo sekä vahvistaa naisten ja tyttöjen oikeuksia ja mahdollisuuksia” ja kahdeksas tavoite on ”edistää kaikkia koskevaa kestävä talouskasvua, täyttä ja tuottavaa työllisyyttä sekä säällisiä työpaikkoja”. Näillä tavoitteilla on alatavoitteita, kuten lopettaa naisiin ja tyttöihin kohdistuva syrjintä ja hyväksikäyttö, turvallisten työolojen turvaaminen sekä työntekijöiden oikeuksien kunnioittaminen henkilön iästä, sukupuolesta tai vammaisuudesta riippumatta. (Ulkoministeriö, n.d.; United Nations Economic and Social Council, 2016, ss. 21, 24)

Syrjinnän ennaltaehkäisytyö kytkeytyy vahvasti edellä mainittuihin kestävän kehityksen tavoitteisiin myös yhdenvertaisuuslain (1325/2014) ja lain naisten ja miesten välisestä tasa-arvosta (306/1986) tarkoitusten kautta.

2 Kehittämistyön taustat

Tämän opinnäytetyön teoreettinen viitekehys koostuu työsyrynnän esiintymistapojen lisäksi tietojohdamisen merkityksestä, datan laatutekijöistä, analytiikan tasoista sekä tilannekuvan määritelmästä. Työhön valittu näkökulma tehostamiseen on raportointiin kuluvan resurssin eli käytettävän ajan vähentäminen eli raportointityön tehostaminen tarkoittaisi työn tekemistä vähemmällä panostuksilla (Eskelinen ym., 2017, ss. 5–6).

Syrjinnän esiintymismuodoista ja määräst, tietojohdamisesta sekä raportointijärjestelmien kehitystyöstä ja prosesseista löytyy kirjallisuutta ja ajankohtaista tutkimustietoa. Tässä osassa syvennyttään valikoidusti näihin aiheisiin.

2.1 Työsyryntä

Yhdenvertaisuuslain (1325/2014) 8 §:ssä säädetään siitä, ettei ketään saa syrjiä ”*iän, alkuperän, kansalaisuuden, kielen, uskonnon, vakaumuksen, mielipiteen, poliittisen toiminnan, ammattiyhdistystoiminnan, perhesuhteiden, terveydentilan, vammaisuuden, seksuaalisen suuntautumisen tai muun henkilöön liittyvän syyn perusteella*”. Syryntä on säädetty rikoslaisa (39/2001) rangaistavaksi teoksi. Lisäksi työopimuslaisa (55/2001) on erityissäännöksiä työsyryntään liittyen.

Työelämän tutkimuskeskuksen ja Tampereen yliopiston toteuttaman työsyrynnän kokemusta koskevan tutkimuksen haastatteluissa oli kerätty tietoa henkilöiden subjektiivisista kokemuksista. Tutkimus osoitti kattavasti sen, minkälaisa syryntä esiintyy suomalaisilla työpaikoilla, mutta siinä ei käsitelty kaikkia syryntäperusteita eikä siitä löytynyt edes anonymisoituna suoria lainauksia haastatteluista. Todennäköisesti tähän ratkaisuun oli päädytty työsyryntään liitettävien negatiivisten miellelyhtymien takia. (Pietiläinen ym., 2018)

Oikeusministeriön vihapuhetta julkaisusta löytyi jonkin verran suoria lainauksia tutkimusta varten toteutetuista haastatteluista sekä kyselyn avovastauksista, mutta nekin edustavat yksittäisten henkilöiden kokemuksia. Tätä tutkimusta varten oli haastateltu eri vähemmistöryhmiä heidän kokemastaan vihapuheesta ja häirinnästä. Tutkimuksessa todettiin, että vihapuhe voi olla kielletyn syrynnän esiintymismuoto. On kuitenkin huomattava se, ettei tutkimusta ole toteutettu työsyryntäilmiön näkökulmasta, vaikka tutkimuksessa tuotiin esiin se, että rikoslain, yhdenvertaisuuslain ja tasa-arvolain lisäksi myös työturvallisuuslaisa (738/2002) säädetään työntekijälle terveyttä haittaavan ja vaarantavan häirinnän kiellosta. (Jauhola ym., 2022, ss. 11, 15, 18) Taulukossa 1 esitetään vihapuhetta koskevasta

tutkimuksesta löytyviä esimerkkejä siitä, miten henkilöt kuvaavat tapahtunutta häirintää (Jauhola ym., 2022, ss. 62–172).

Taulukko 1. Kuvauksia häirinnästä (Jauhola ym., 2022, ss. 62–172)

<i>Vähemmistöryhmä</i>	<i>Kuvauksia henkilöiden kokemasta häirinnästä</i>
<i>Vammaisuus</i>	<i>tuijotus</i> , naureskelu, osoittelu, <i>nimittely</i> , nolaaminen, nöyryyttäminen, <i>huutaminen</i> , supistaan ikäviä kommentteja, ei uskota, <i>vähätellään</i> , kyseenalaistetaan, rajoittaminen, avun epääminen, liika koskettelu, intiimiasioden kertominen kovaäänisesti, työn kuvan muuttamien toisten toiveesta, etujen epääminen
<i>Romanit</i>	perässä kulkeminen, seuraaminen, <i>huutelu</i> , nöyryyttäminen, haukkuu, palvelee huonommin, ivaaminen, uhkailu, nimittely, osoittelu, alentavia nimityksiä
<i>Vieraskieliset, ulkomaan kansalaiset, maahanmuuttajataustaiset</i>	rasistinen nimittely, <i>huuhtelua</i> , pahasti katsominen tai vihamielinen <i>tuijottelu</i> , <i>ulosulkeminen</i> , <i>vähättelyä</i> , <i>väheksyminen</i> , terroristiksi kutsuminen, kiusaamista, fyysisesti käsiksi käymistä, halventamista, rodullistamista
<i>Seksuaali- ja sukupuolivähemmistöt</i>	häiritsevä <i>huutelu</i> , ivaaminen, alistaminen, homottelu, transuttelu, transfobinen häirintä, transfobinen utelu, väärinsukupuolittaminen, <i>uhkaava</i> käytös, fyysinen väkivalta, <i>uhkaavat</i> käsimerkit, <i>haukkuminen</i> julkisella paikalla, käskyt tappaa itsensä, <i>väheksyntä</i> , <i>vähättely</i> , <i>tuijotus</i> , mulkoilu, provosoiminen, huomiotta jättäminen, <i>ulkopuolelle sulkeminen</i> , ei suostuta käyttämään uutta nimeä, loukkaavien termien käyttö, määräaikaisen työsopimuksen uusimisen peruuttaminen, vitsailu työpaikalla, pilkkaaminen, dissaaminen,
<i>Uskonnolliset vähemmistöryhmät</i>	<i>nimittely</i> (terroristiksi, lampaaksi, typeräksi, synnytyskoneeksi), <i>haukkuminen</i> (terroristiksi, raiskaajien kasvattajaksi, kummitukseksi), ”rätti pois päästä”, töniminen, <i>huutelu</i> , ilkeily, selän takana puhuminen, syljeskelyä, naureskelua, <i>vähättelyä</i> , huomiotta jättämistä, tuhahtelua, silmien pyörittelyä, <i>ulkopuolelle jättämistä</i> , uskonnon haukkumista, takaapäin puhumista, arvostelua, rikosten yleistäminen koko uskontokuntaan
<i>Saamelaiset</i>	<i>haukkumista</i> , halveksuntaa, kyseenalaistamista, kiusantekoa, vaientamista, asiattomuuksien utelu, alistamista, elinkeinoon puuttamista, negatiivinen keskustelu, <i>huutelua</i> , osoittelua, naurun alainen, <i>vähättelyä</i> , rasistiset ilmaisut, väkivallalla <i>uhkailu</i> , perättömien rikosilmoitusten teko, <i>uhkaava</i> käytös, <i>huutaminen</i> , sivuun työntäminen, koskettelu työpaikalla, vääristelyä, disinformaatiota, nälvimistä
<i>Uskontokuntaan kuulumattomat</i>	syrjimistä, <i>nimittely</i> (valehtelijaksi, ihmisvihaajaksi, vainoajaksi), selän takana <i>haukkuminen</i> , osoittelu, leimaaminen, uskonnottomuuden ohittaminen, <i>uhkailu</i> , moraalittomaksi <i>haukkuminen</i>

Taulukosta 1 nähdään, että eri häirintää koskevat ilmenemistavat ovat hyvin samankaltaisia vähemmistöryhmien kesken. Eri vähemmistöryhmät kokevat hyvin usein muun muassa tuijottelua, huutelua, haukkumista, nimittelyä, vähättelyä, ulkopuolelle jättämistä ja uhkaavaa käytöstä.

Vihapuhetta koskevassa tutkimuksessa todettiin, että vastaajat kokivat työelämässä vihapuhetta tai häirintää sekä työkavereiden että esihenkilöiden suunnasta. Tutkimukseen osallistuneista maahanmuuttotautaisista 20 % oli kokenut vihapuhetta työkaverilta. Vastaava luku romaneilla oli 14 %, saamelaisilla 15 %, vammaisilla 16 %, seksuaali- ja sukupuolivähemmistöihin kuuluvilla 14 % sekä uskonnollisiin vähemmistöihin kuuluvilla 26 %. Esihenkilön suunnasta vihapuhetta oli kokenut maahanmuuttotautaisista vastaajista 7 %, saamelaisista 10 %, vammaisista 12 %, seksuaali- ja sukupuolivähemmistöihin kuuluvista 5 % ja uskonnollisiin vähemmistöihin kuuluvista 2 %. (Jauhola ym., 2022, s. 41)

Työsyrynnästä löytyi kattavasti muutakin tutkimustietoa, muun muassa Pietiläisen ym. (2018) laatima loppuraportti työssä koetusta syrjinnästä ja sen vaikutuksista myöhempään työuraan sekä Valtioneuvoston kanslian julkaisu työsyrynnän nykytilasta ja keinoista tasa-arvon ja yhdenvertaisuuden edistämiseksi (Kanninen ym., 2022). Olemassa olevasta tutkimustiedosta ei kuitenkaan löydy riittävästi sellaista tietoa, minkä perusteella pystyttäisiin muodostamaan toimivia hakusanoja tai sanojen yhdistelmiä asiakasyhteydenottojen sisällön tilannekuvan luontia varten. Tämän takia tässä opinnäytetyössä päädyttiin analysoimaan työsuojeluviranomaiselle tehtyjen asiakasyhteydenottojen kuvaustekstejä.

Vuonna 2009 sisäasiainministeriön julkaisemassa tutkimuksessa oli hyödynnetty työsuojelupiirien aineistoa, joka sisältää henkilöiden subjektiivisia kokemuksia työsyrynnästä. Tähän tutkimusaineistoon oli kuulunut vain ne asiakasyhteydenotot, joiden perusteella viranomaiset olivat toteuttaneet valvontatoimenpiteitä. Tutkimusaineiston rajoittamiseen oli vaikuttanut aikataulupaineiden lisäksi se, ettei kiellettyä syrjintäperustetta pystytty yksilöimään asiakkaan yhteydenotosta. (Aaltonen ym., 2009, ss. 21–22)

Samassa sisäasiainministeriön julkaisemassa tutkimuksessa työsuojeluviranomaiselle esitettiin kehitysehdotuksena, että tapaustietoihin tulisi kirjata myös syrjintää kokeneen henkilön ikä ja sukupuoli, koska näiden tietojen avulla olisi mahdollista seurata moniperusteisen syrjinnän esiintymistä. Kehitysehdotuksena esitettiin myös tapaustietojen kirjaamisen yhdenmukaistamista sekä kirjausten päivittämistä asian käsittelyn edetessä. Näiden lisäksi esitettiin lisättäväksi tapauksen tietoihin myös tieto mahdollisesta syrjintäperusteesta. Kehitysehdotusten yhteydessä oli todettu, ettei työsuojeluviranomaisen asianhallintajärjestelmä tue tutkimustarpeita. (Aaltonen ym., 2009, s. 26)

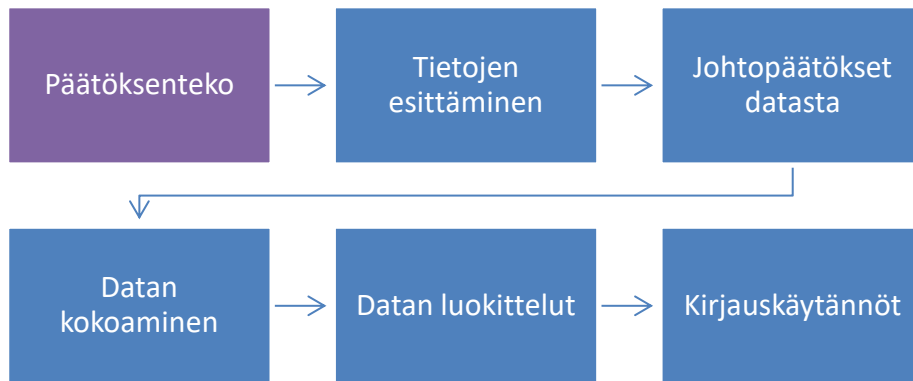
Tässä opinnäytetyössä palataan saman tietolähteen äärelle kuin mitä oli hyödynnetty sisäasiainministeriön julkaisemassa tutkimuksessa eli asiakasyhteydenottojen kuvauksiin. Asiakasyhteydenottojen kuvaukset oli jätetty aiemman tutkimusaineiston ulkopuolelle. Työsuojeluhallinnossa on tapahtunut monia muutoksia tutkimuksen julkaisun jälkeen, joista ehkä merkittävin on se, että työsuojeluviranomainen on ottanut uuden valvontatietojärjestelmä käyttöön vaiheittain vuosien 2011–2014 aikana. Uuden järjestelmän käyttöönoton jälkeen asiakasyhteydenottojen kuvaukset löytyvät valvontatietojärjestelmästä (Sosiaali- ja terveystieteiden ministeriö, 2016, ss. 47–49, 55, 98).

2.2 Tietojohtaminen

Vuonna 2009 sisäasiainministeriön julkaisemassa tutkimuksessa oli tehty havainto siitä, että työsuojeluviranomaisen asianhallintajärjestelmä ei tue tutkimustarpeita. Tilanne on täysin ymmärrettäviä työsuojeluviranomaisen ydintehtävän näkökulmasta. Aluehallintovirastojen työsuojelun vastuualueet eivät ole tutkimusorganisaatioita, vaan ne suorittavat niille annettua lakisääteistä valvontatehtävää (Laki työsuojelun valvonnasta ja työpaikan työsuojeluyhteistoiminnasta 44/2006). Tästä huolimatta kehitysehdotus syrjintäperusteen kirjaamisesta on edelleen relevantti tietojohtamisen näkökulmasta.

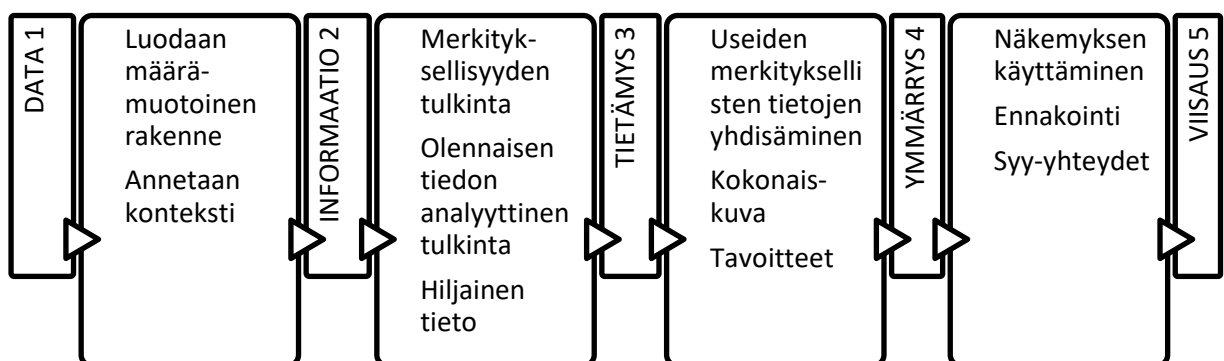
Tietojohtamisen kattokäsite jakautuu kahteen kokonaisuuteen – tiedon johtamiseen sekä tiedolla johtamiseen. Tietojohtamisen tekniset näkökulmat datan hankinnasta, varastoinnista, siirtämisestä ja suojaamisesta sekä kirjaamiskäytännöistä tietojärjestelmiin muodostavat *tiedon johtamisen* kokonaisuuden. Organisaation datasta jalostetun informaation perusteella tehtyjen johtopäätösten esittäminen sekä tiedon hyödyntäminen päätöksenteossa muodostaa *tiedolla johtamisen* kokonaisuuden. (Leskelä ym., 2019, s. 16; Listenmaa, 2023, ss. 45–46) Datasta jalostettu tieto auttaa organisaatioita tekemään perustellumpia jopa parempia ratkaisuja niin strategisella, operatiivisella kuin taktisella tasolla, jotta organisaatio saavuttaa toimintatavoitteensa (Listenmaa, 2023, s. 123). Ideaalitalanteessa (kuva 1) tiedolla johtamisen kokonaisuus ohjaa tiedon johtamista eli organisaation strategiset toimintatavoitteet sekä päätöksenteon tietotarpeet ohjaavat tiedon keräämisen suunnitelmaa (Leskelä ym., 2019, ss. 50–51).

Kuva 1. Tietojohtamisessa päätöksenteko ohjaa tiedon keräämistä ja analyysiä (Leskelä ym., 2019, s. 50)



Yleisenä käsitteenä tiedolla tarkoitetaan merkkijonoa, havaintoa tai tosiasiaa (Finto Suomalainen asiasanasto- ja ontologiapalvelu, n.d.b). Tietojohtamisen viitekehyksessä tiedolle on määritetty eri tasoja eli jalostusasteita. Jalostustasojen määrä vaihtelee eri julkaisuissa. Laihonen ym. (2013, s. 18) mainitsevat kolme tasoa, jotka ovat *data*, *informaatio*, *tietämys*. Cong & Pandya (2003, ss. 26–27) käyttävät neljää tasoa, jotka ovat *data*, *information*, *knowledge*, *wisdom*. Listenmaa (2023, s. 36) yksilöi viisi tasoa, jotka ovat *data*, *informaatio*, *tieto*, *näkemys*, *viisaus*. Kaikkia näitä yhdistää hierarkkinen eteneminen. Lisäksi yhteisiltä luokituksiltaan ja tiedon jalostumisen osalta mainittuja jalostusasteet eivät ole keskenään ristiriidassa. Olennaista tiedon jalostusasteissa on kuitenkin se, mitä tapahtuu tasolta toiselle siirtyessä, jota on kuvattu kuvassa 2.

Kuva 2. Tiedon jalostuminen siirtymissä. (Cong & Pandya, 2003, s. 26; Finto Suomalainen asiasanasto- ja ontologiapalvelu, n.d.b; Listenmaa, 2023, ss. 35–38)



Data on tosiasioita, joita ei välttämättä ole esitetty määrämuotoisesti. Data on yksittäisiä havaintoja, jotka sellaisenaan eivät tuota mitään lisäarvoa. Datan jatkokäytettävyyden kannalta datan tulee olla laadukasta, virheetöntä ja paikkansa pitävää, jotta sitä pystytään täysimääräisesti hyödyntämään. Kuvasta 2 nähdään, että irralliset, yksittäiset merkkijonot tai

havainnot muuntuvat informaatioksi, kun niille luodaan rakenne. (Sivula ym., 2023; s. 60; Listenmaa, 2023, s. 46; Laihonen ym. 2013, ss.16–17)

Informaatiolla tarkoitetaan analyyseissä hyödynnettävää rakenteisesti esitettyä dataa ja tiedolla informaatiota, jota on rikastettu kokemukseen perustuvalla inhimillisellä tiedolla. Kuvasta 2 nähdään myös, että olennainen ja merkityksellinen informaatio kumuloituu ajan saatossa ja analysoinnin kautta muuntuu tietämykseksi. Tietämys tarkoittaa siis merkityksellistä informaatiota. Informaation muuntumiseen tietämykseksi liittyy myös hiljaisen tiedon hyödyntäminen tulkinnassa. Hiljainen tieto on ihmisten kokemusperäistä ja oppimisen kautta muodostunutta tietämystä (Listenmaa, 2023, s. 26). Ymmärrys muodostuu asiayhteyteen liittyvien merkityksellisten tietojen kokonaiskuvasta ja niiden kytköksistä toiminnan tavoitteisiin. (Cong & Pandya, 2003, s. 26; Finto Suomalainen asiansanasto- ja ontologiapalvelu, n.d.b; Laihonen ym., 2013, ss. 16–17; Listenmaa, 2023, ss. 35–38, 46; Sivula ym., 2023, s. 60)

Organisaation hallussa oleva tieto nähdään nykyisin samalla tavalla hyödynnettävänä resurssina kuin mitkä tahansa muutkin työvälineet (Laihonen ym., 2013, s. 10). Työsuojeluviranomaista voi pitää asiantuntijaorganisaationa, jolloin sen tiedolla johtamisen strategia painottuisi enemmän Listenmaan (2023, ss. 117–120) kuvaamaan ihmiskeskeiseen *personointistrategiaan* kuin tekniseen kodifointistrategiaan. Personointistrategiassa korostuvat vuorovaikutuksessa tapahtuva tiedon luonti, hyödyntämien sekä levittäminen. Työsuojeluviranomaisen hallussa olevan datan analysoinnilla ei niinkään ole suurta merkitystä organisaatiolle syntyvän kilpailuedun saavuttamisessa tai uusien liiketoimintamahdollisuuksien löytämisessä, koska viranomaisen toteuttaa sille annettua lakisääteistä tehtävää. Suurin merkitys onkin työsyöjinnän minkä tahansa muun valvottavan ilmiön eri esiintymismuotojen havainnoinnissa, näiden trendien suunnanmuutosten tunnistamisessa sekä kertyneen tiedon hyödyntämisessä viranomaisen toiminnan suuntaamisessa ja päätöksenteossa. (Listenmaa, 2023, ss. 123–125; Takala, 2018, s. 3)

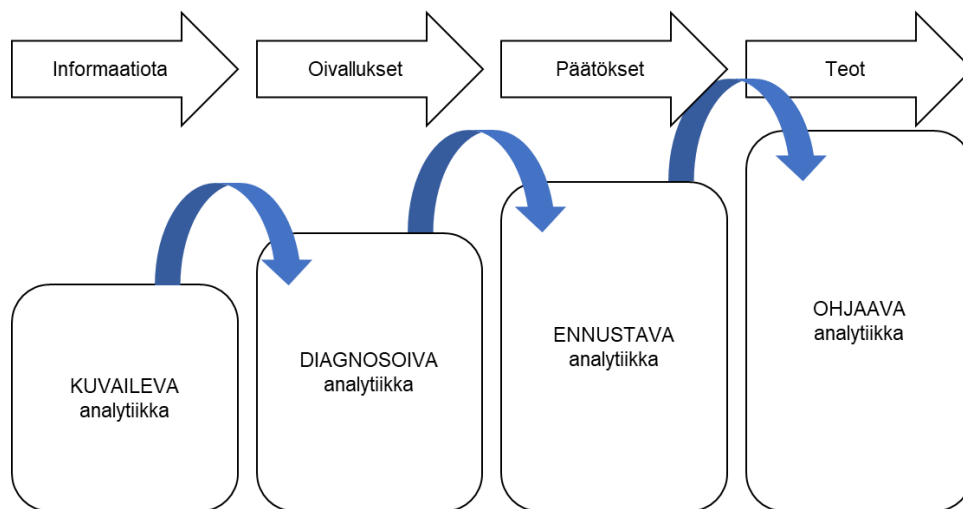
2.3 Raportointi ja analytiikka

Raportointi on vakiintunut osa julkishallinnon tehtäviä. Raportoinnilla on roolinsa tiedolla johtamisen ja toiminnan kehittämisen näkökulmasta niin julkishallinnossa raportoinnilla todennetaan muun muassa tulosohjauksen toteutumista. (Listenmaa, 2023, ss. 231–234; Salminen ym., 2021)

Toteutunutta toimintaa kuvaava raportointi on kuvaileva analytiikka (*descriptive analytics*), jonka tarkoitus on kuvata historiatiedosta sitä, mitä dataa on kerätty tai mitä on toteutettu.

Kuvailevassa analytiikassa voidaan hyödyntää historiatietoa laajastikin. Muut vakiintuneet analytiikan suuntaukset ovat: diagnosoiva, ennustava sekä ohjaava analytiikka. Nämä neljä analytiikan tasoa ovat keskenään hierarkkisia, joka tarkoittaa samalla sitä, että analysoinnin vaikeustaso nousee aina siirryttäessä seuraavalle tasolle (kuva 3). (Listenmaa, 2023, ss. 122, 275–277; Sivarajah ym., 2017, ss. 266, 275–277)

Kuva 3. Analytiikan tasot Sivarajah ym. (2017, s. 226) mukaillen



Diagnosoiva analytiikka (*Inquisitive analytics*), etsii eri tekijöiden yhteyksiä ja vaikutuksia yhdistämällä tietoja tai hyödyntämällä tiedonlouhintamenetelmiä kuten tiivistämistä, klusterointia taikka esimerkiksi aikasarjan luontia (Ngai ym., 2009, s. 2593). Ennakoivan analytiikan (*predictive analytics*), tarkoitus on tuottaa näkemystä potentiaalisista tapahtumista tai trendien kehityssuunnista. Ohjaava analytiikka (*prescriptive analytics*), tavoittelee toiminnan optimointia tai simuloi muutostilanteita. (Listenmaa, 2023, ss. 122, 275–277; Sivarajah ym., 2017, s. 266; Takala, 2018, ss. 5–6).

2.4 Tilannekuva

Tilannekuva on ”koottu kuvaus vallitsevista olosuhteista, käsillä olevan tilanteen synnyttäneistä tapahtumista, tilannetta koskevista taustatiedoista ja tilanteen kehittymistä koskevista arvioista sekä eri toimijoiden toimintavalmiuksista” (Sanastokeskus, 2023). Tilannekuvan muodostaminen vaatii siten toimintaympäristön havainnointia ja näiden havaintojen tallentamista dataksi. Toimintaympäristön havainnointi tarkoittaa tietoisuutta vallinneista olosuhteista sekä olosuhteiden taustatiedoista. Tilannekuvassa kerätty data on jalostettu informaatioksi ja tietämykseksi. (Listenmaa, 2023, ss. 122, 275–277; Mykkänen ym., 2021, s. 426)

Tilannekuvan sisältö voi muistuttaa reaaliaikaista raportointia ja näiden kahden välille on vaikea tehdä selkeää rajanvetoa. Tilannekuvan selkeä rooli on kuitenkin auttaa ymmärtämään tapahtumien nykytilaa sekä siihen vaikuttaneita asioita, jotta organisaatio voisi ennakoida parasta seuraavaa toimintatapaa. Tilannekuva auttaa siten organisaatiota päätöksenteossa ja siksi tilannekuvan tulee sisältää organisaation tavoitteista johdettuja, analysoituja sekä jäsenneltyä tietoa. Tilannekuvan merkitys organisaatioille liittyy usein toiminnan seurantaan ja ohjaukseen ja siinä olevat tunnusluvut tukevat päätöksentekoa. Toimiva tilannekuva tukee organisaation vision ja tavoitteiden toteutumista (Cahyadi & Prananto, 2015, ss. 256–290; Listenmaa, 2023, ss. 122, 275–277; Mykkänen ym., 2021, s. 426)

Tilannekuvalta odotetaan usein jonkinlaista visuaalista esitystapaa organisaation toiminnan tilanteesta eikä niinkään tekstiksi koottua kokonaiskuvaa. Visuaalisuuden lisäksi tilannekuvassa tulisi olla myös jonkin verran toiminnallisuuksia, esimerkiksi mahdollisuus porautua yksityiskohtaisempaan tiedon tasoon tai näkymän osittamista. (Cahyadi & Prananto, 2015, ss. 286–303; Sivarajah ym., 2017, s. 275)

Tilannekuva kertoo siis nykytilasta ja johtamisen työvälineenä sen avulla voidaan ennakoida lyhyellä tai pitkällä aikavälillä tapahtuvia muutoksia. Tilannekuvan on perustuttava relevanttiin tilannetietoon, joka vaatii sujuvuutta datan keräämisessä sekä tiedon siirtymisessä organisaatiossa. (Kuusisto, 2005, ss. 8–9; Mykkänen ym., 2021, s. 426)

2.5 Datan laatu automatisoinnin edellytyksenä

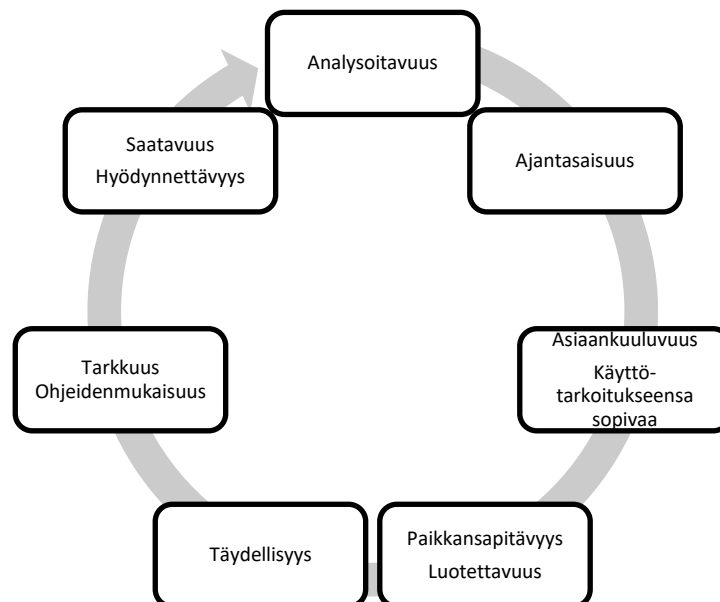
Toiminnan tehostamisen toteuttamistavoista ehkä yleisin toteutustapa on toistuvien ja rutiiniluontoisten tehtävien automatisointi. Yleisesti automatisoinnin toteutuksen edellytys on, että data on luotettavaa ja laadukasta. Tämän lisäksi löytyy joukko muita teknisempiä edellytyksiä erityisesti, jos hyödynnetään tekoälyä. Teknisempiin edellytyksiin kuuluvat niin tiedon omistaminen kuin tiedon tekninen rakenne ja esittämismuodot. (Aihkisalo ym., 2018, ss. 7, 21–24; Sivula ym., 2023, ss. 42–43)

Automatisoinnin edellytyksenä datan korkean laadun lisäksi voidaan pitää sitä, että data on rakenteellisessa muodossa. Datan rakenteellisuus on olennaista automatisoinnille, koska rakenne tuottaa metatiedot, joita automatisoinnissa hyödynnetään. Metatiedot tarkoittavat tietoa, joka antaa datalle kontekstin eli määrämuotoisen sisällön. Rakenteettomalta datalta puuttuu metatiedot, joten sitä on vaikea käsitellä koneellisesti. Rakeellisessa muodossa oleva data mahdollistaa tehokkaan raportoinnin, koska tiedot ovat poimittavissa suoraan tietokannasta niin usein kuin on tarve. Lisäksi datan rakenteellisuus johtaa siihen, että data

voidaan tallentaa yhden kerran, mikä vähentää datan manuaalista käsittelyä. (Finto Suomalainen asiasanasto- ja ontologiapalvelu, n.d.a, n.d.c.; Gil ym., 2022; Mykkänen ym., 2021, s. 427, 2023, s. 151; Valtionvaraniministeriö, 2019).

Yleisesti laadukkaalla datalla on tarkoitettu käyttötarkoitukseensa sopivaa ja tarvittavat muuttujat sisältävää dataa (Väre, 2019, ss. 198–199). Tarkemmin arvioituna laadukas data on analysoitavissa, ajantasaista ja asiaankuuluvaa eli relevanttia käyttötarkoitukseensa nähden. Lisäksi laadukas data on saatavissa, täydellistä, tarkkaa sekä paikkansapitävää eli luotettavaa (kuva 4) (Sivula ym., 2023, s. 60).

Kuva 4. Datan laatutekijät mukailten Sivula ym. (2023, s. 60) ja standardi SFS-ISO/IEC 25012:2020



Ajantasaisuus tarkoittaa tilanteeseen nähden oikeaan aikaan hankittua ja tuotettua tietoa. Asiaankuuluvuus tarkoittaa tapahtuman tai käsitteen kannalta käyttötarkoitukseen sopivan datan olemassaoloa. Saatavuus tarkoittaa mahdollisuutta datan lataamiseen ja siirtämiseen muuhun käyttöön tietojärjestelmän ulkopuolelle. Saatavuuteen liittyy myös vaatimus luottamuksellisuudesta eli data on vain niiden käytössä, joilla on siihen oikeus. Tarkkuuteen liittyy vaatimus ohjeidenmukaisuudesta eli datan tulee noudattaa sille määritettyjä ohjeita tai standardeja. (SFS-ISO/IEC 25012:2020, 2020, ss. 11–13; Sivula ym., 2023, s. 44)

Datassa esiintyvät laatupoikkeamat voivat johtua inhimillisistä virheistä datan kirjausvaiheessa, jolloin jotain jää tallentamatta taikka tapahtuu kirjausvirhe. Toinen todennäköinen laatupoikkeama paljastuu raportointivaiheessa, kun havaitaan, ettei data vastaa tietotarvetta. (Listenmaa, 2023, ss. 119, 252) Vaikka data olisi laadukasta niin datan

hyödyntämiseen voi liittyä muita haasteita, kuten strategian tai osaamisen puuttuminen, datan säilytysajan pituus, datan määrä tai jopa muut tietotekniset asiat kuten riittämätön tiedonsiirtonopeus. (Sivula ym., 2023, ss. 39–48).

3 Tarkoitus ja tavoitteet

Toimeksiantaja on todennut työsyynnän vuosiraportointiin vievän merkittävän paljon aikaa, koska datan analysointityössä ei ole käytössä vakioitua raportointisovellusta tai muuta ratkaisua. Toimeksiantaja on päivittäisessä työssä todennut, ettei työsyynnäilmiöön liittyvien asiakasyhteydenottojen määrä riitä kertomaan työsyynnän esiintymistavoista eli syrjintäperusteista. Kansalaiset eivät aina erota sitä, miten kielletty syrjintä ja tasapuolisen kohtelun vastainen toiminta eroavat toisistaan. Lisäksi työsyynnä sekoitetaan monesti häirintään tai epäasialliseen kohteluun. Näiden syiden takia pelkkä yhteydenottomäärän raportointi ei tuota riittävän luotettavaa ja todenmukaista kokonaiskuvaa työsyynnäilmiöstä.

Tämän opinnäytetyön ennakkoselvitysvaiheessa selvisi, että nykyinen raportointitapa vaatii kohtuuttoman paljon manuaalista työtä. Ongelman syyksi tunnistettiin, että yhteydenoton kirjaamisen yhteydessä tietojärjestelmään jää usein tallentamatta tarvittavat lisätiedot. Tämä johtuu siitä, että lisätietojen kirjaaminen on osin muistinvaraista eikä tietojärjestelmä pakota lisätietojen tallentamiseen. Toinen tunnistettu syy tilanteelle on, ettei kaikkea asiakasyhteydenotoista tallennettavaa tietoa tallenneta rakenteellisena datana.

Tieto työsyynnän esiintymismuodosta kirjataan järjestelmään vapaana tekstinä eli muodossa, jolla ei ole ennalta määritettyä tunnistettavaa järjestystä tai sisältöä. Työsyynnäperusteiden analysoinnin kannalta tärkein data on siis rakenteetonta eli se on sellaisessa muodossa, jossa sen hyödyntämismahdollisuudet ovat alhaisimmat (Finto Suomalainen asiasanasto- ja ontologiapalvelu, n.d.a). Edellä mainituista syistä syrjintäperusteiden analysointi on täysin manuaalista työtä eli vaatii jokaisen yhteydenoton kuvauksen lukemisen ja johtopäätöksen tekemisen. Toimeksiantajalla ei ole tällaiseen työhön riittäviä resursseja, joten asiakasyhteydenotoista ilmeneviä syrjintäperusteita on raportoitu vain esimerkinomaisesti vuosiraportoinnissa.

Toimeksiantajan tarve kehitystyölle liittyy nimenomaisesti raportoinnin tehostamiseen ja nopeuttamiseen sekä syrjintäperusteiden tunnistamiseen. Työn päätavoite on parantaa toimeksiantajan mahdollisuuksia julkaista ajantasaisempaa tilastotietoa suomalaisen työelämän työsyynnän esiintymismuodoista. Analysointityön tehostamisen konkreettiseksi kehitysteoksi valikoitui raportointityökalun luonti. Tavoitteena oli luoda toimeksiantajalle raportointityökalu, jolla voidaan tuottaa visuaalinen tilannekuva asiakasyhteydenotoista.

Raportointityökalun ja tilannekuvan päivittämisen tulisi tapahtua mahdollisimman vähän toimenpiteitä.

Syrjintäperusteita koskevan tilannekuvan kannalta tärkein data on alkutilanteessa rakenteetonta, jolloin sen hyödyntämismahdollisuudet ovat alhaisimmat. Työn tavoitteena on lisätä rakennetta asiakasyhteydenoton rakenteettomaan tekstidataan ja saadaan data paremmin hyödynnettävään muotoon. Raportointityökalun kehittämisen lopputuotteena valmistuu myös työohje toimeksiantajan käyttöön.

Raportoinnin kehittämisestä hyötyvät työsuojeluhallinto ja sen sidosryhmät, kun syrjinnän muodoista saadaan helpommin ja nopeammin muodostettua kokonaiskuva. Lisähyötyä saavutetaan työsuojeluhallinnossa laajemminkin, jos toteutettava raportointityökalu voidaan monistaa muihin asiakasyhteydenottojen asialuokkiin. Toimeksiantaja näkee mahdollisena myös sen, että tilannekuvaa pystyttäisiin hyödyntämään myös viestinnän kohdentamisessa.

Työn uutuusarvo perustuu erityisesti siihen, että työsuojeluviranomaiselle tehtyjen asiakasyhteydenottojen sisältöä analysoidaan aiempaa tarkemmin ja tällä tavalla tuotetaan uutta tietoa henkilön eli asiakkaan syrjinnän kokemuksesta subjektiivisesta näkökulmasta. Lisäksi analysoinnin voidaan katsoa tehdyksi ulkoisena arviointityönä, kun opinnäytetyön toteuttaja ei ole osallistunut työsyrijintään koskevien asiakasyhteydenottojen käsittelyyn, ohjaukseen eikä syrjintätapausten valvontaan missään roolissa (Opetushallitus, 1998, s.14). Tähän asti syrjintään liittyvä analysointityö on tehty viranomaisen sisäisenä arviointityönä.

4 Tutkimusmenetelmät

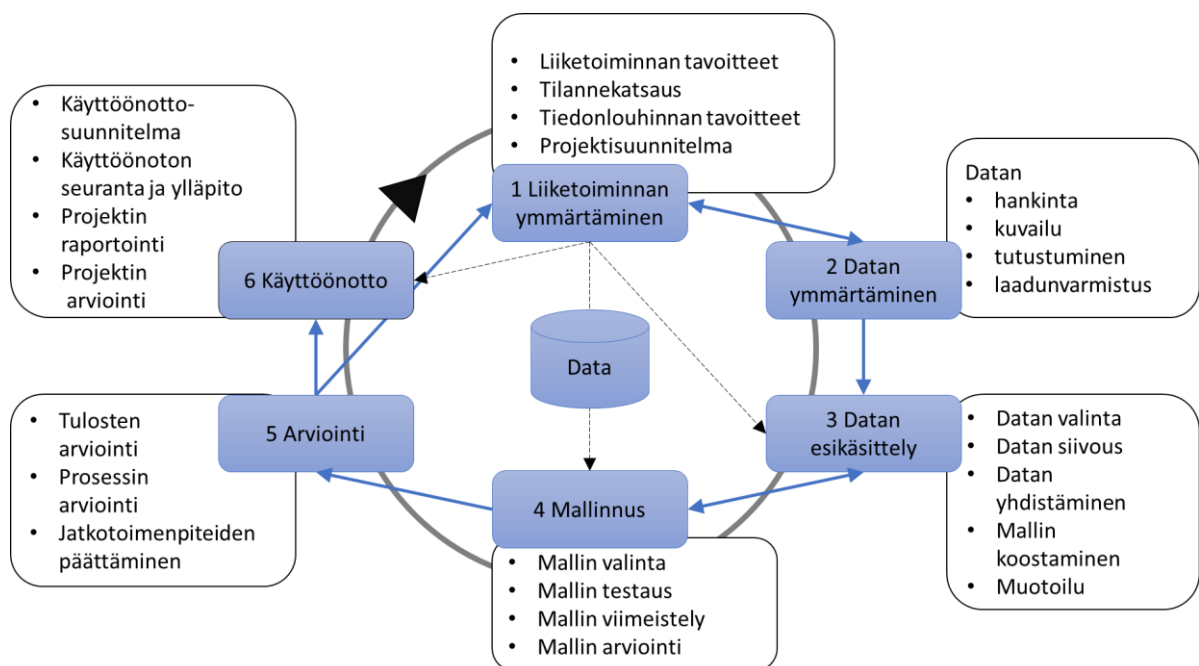
Työelämään liittyvä kehittäminen tapahtuu usein toimintatutkimuksen periaatteiden mukaisesti. Toimintatutkimuksen vakiintuneet vaiheet ovat suunnitelman laatiminen, toiminnan toteutus, havainnointi, toteutuksen reflektointi sekä suunnitelman parantaminen. Kehittäminen jatkuu spiraalimaisesti näitä vaiheita toistaen. (Jyrkämä, n.d.) Nämä vaiheet toteutuvat myös CRISP-DM- tiedonlouhintaprosessissa, mikä valittiin tämän kehittämistyön menetelmäksi.

Opinnäytetyön tutkimusaineisto ja erityisesti asiakasyhteydenoton kuvaus on tekstiä, ja kehittämistyön tavoitteet edellyttävät tutkimukselta vahvaa aineistolähtöistä lähestymistapaa. Asiakasyhteydenottojen sisällönanalysointi toteutetaan teoriapohjaisesti temaattisena analyysinä. Toimeksiantajalle luotavan tilannekuvan määrittelyt kerättiin puolistrukturoituna ryhmähaastatteluna. Raportoinnin ja tilannekuvan kehittämisessä hyödynnetään viranomaisen valmista aineistoa eli järjestelmään tehtyjä kirjauksia.

4.1 Tiedonlouhintaprosessi CRISP- DM

Viimeisen 20 vuoden aikana tiedonlouhinnan (*data mining*) sekä analysoinnin vakiintuneeksi prosessimalliksi on muodostunut CRISP-DM (myöhemmin CRISP). Kuusivaiheista CRISP-tiedonlouhintaprosessia on hyödynnetty monilla eri toimialoilla, myös kirjallisuuskatsauksien ja tekstiaineistojen tiedonlouhinnassa. Kuvassa 5 on esitetty CRISP:n vaiheet, jotka ovat: liiketoiminnan tarpeiden ymmärtäminen, dataan tutustuminen ja ymmärtäminen, datan esikäsittely eli siivous ja jalostaminen, mallintaminen, arviointi ja käyttöönotto. Jokainen vaihe koostuu useasta osakokonaisuudesta, jotka on myös esitetty kuvassa 5. (Jackson, 2002, s. 283; Martinez-Plumed ym., 2021, ss. 1–3; Schröer ym., 2021, ss. 526–529).

Kuva 5. CRISP-DM-prosessin vaiheet laajennetuin tiedon sekä ensimmäisen vaiheen riippuvuussuhteet muihin vaiheisiin (Jackson, 2002, s. 283; Sharma & Osei-Bryson, 2009, s. 4115).



CRISP on erittäin prosessi- ja tavoitekeskeinen, koska sen lähtökohdat nojaavat vahvasti liiketoiminnan tarpeiden ymmärtämiseen ja analysoinnilla pyritään kuvaamaan juuri ennalta määritettyjä tietoja olemassa olevasta datasta (Jackson, 2002, s. 283; Martinez-Plumed ym., 2021, ss. 1–3; Schröer ym., 2021, ss. 526–529). Ensimmäisen vaiheen ohittaminen kokonaan tai riittämätön syventyminen liiketoiminnan tarpeisiin voi johtaa kehitystyön tehottomuuteen tai työ voi edetä täysin eri suuntaan kuin on tarkoitettu. Näiden riskien takia liiketoiminnan tarpeiden ymmärtämiseen kannattaa käyttää riittävästi resursseja kehitystyön alkuvaiheessa. Liiketoiminnan tarpeet määrittävät analysoinnissa tarvittavan datan,

ratkaisevat mallintamisessa käytettäviä metodeja, yksilöi arviointikriteereitä sekä auttaa ratkaisujen käyttöönotossa. (Sharma & Osei-Bryson, 2009, s. 4115).

Liiketoiminnan tarpeiden ymmärtämistä varten on tehtävä perusteellinen kartoitus nykytilasta sekä toiminnan kehittämisen tarpeista ja tavoitteista ja niihin liittyvistä odotuksista ja rajoitteista samalla, kun keskenään ristiriitaisia tavoitteita priorisoidaan tai muilla tavoin yhteensovitetään. Tässä yhteydessä määritellään myös tiedonlouhintatavoitteet, jotka tukevat liiketoiminnan kehittämisen tavoitteita sekä laaditaan projektisuunnitelma. Hyvin toteutettu liiketoiminnan tarpeiden tunnistaminen voi paljastaa riippuvuussuhteita muihin CRISP-prosessin vaiheisiin, esimerkiksi datan esikäsittelevävaiheeseen, mallinnukseen sekä käyttöönottovaiheeseen (kuva 5). Näitä riippuvuussuhteita voidaan selvittää mikä, miksi, milloin, kuka ja kuinka -kysymyksillä. (Jackson, 2002, s. 282; Sharma & Osei-Bryson, 2009, ss. 4115–4117)

Toisessa vaiheessa hankitaan kokonaiskuva datasta. Vaihe koostuu datan lataamisesta ja erityisesti datan laadunvarmistuksesta. Vaiheen aikana datan perusteella voidaan tehdä ennako-oletuksia. Kolmannessa vaiheessa dataa esikäsitellään eli raakadataa jalostetaan tai puhdistetaan sellaiseen muotoon ja sisältöön, että dataa voidaan hyödyntää mallintamisessa. Jalostamisella tarkoitetaan esimerkiksi uusien muuttujien laskentaa olemassa olevien tietojen perusteella. Puhdistamisella tarkoitetaan sitä, että varmistetaan, että mallinnuksessa käytetään täydellistä datasettiä eli muodoltaan virheelliset arvot on korjattu taikka puuttuvat tyhjät arvot on oikein merkitty aineistoon. (Jackson, 2002, ss. 284–285)

Tietopohjaisten ja parempien ratkaisujen tekemistä auttaa ennustavat analyysit. Tätä tarkoitusta varten neljännessä vaiheessa dataa mallinnetaan eri metodein, kuten regressioanalyysillä, luokittelemalla taikka tunnistamalla toistuvia kaavoja (Ngai ym., 2009, s. 2593). Olennaista on, että mallit soveltuvat liiketoimintatavoitteiden saavuttamisen tueksi. (Jackson, 2002, ss. 285–286)

Ennen käyttöönottovaihetta arvioidaan sitä, että onko liiketoiminnan kannalta olennaiset tavoitteet sekä tiedonlouhinnalle asetetut tavoitteet saavutettu. Kun tavoitteet on riittävässä määrin saavutettu, päätetään siitä, edetäänkö käyttöönottoon vai tarkennetaanko luotua mallia. Käyttöönottovaiheeseen kuuluu myös seuranta- ja ylläpitovaihe. (Jackson, 2002, s. 289)

4.2 Teoriapohjainen temaattinen analyysi

Tilannekuvan muodostaminen asiakasyhteydenottojen sisällöstä edellyttää kehittämiseen liittyvältä tutkimukselta aineistolähtöistä lähestymistapaa. Tämän työn tavoitteena ei ole uuden teorian muodostaminen eikä tulosten laajempi yleistettävyyys vaan syvällinen näkemys aineistossa esiintyvistä tunnistettavista syrjinnän muodoista. Tilannekuvan muodostaminen edellyttää, että järjestelmään tehdyistä kirjauksista tunnistetaan se, mitkä kaikki syrjintäperusteet esiintyvät järjestelmään tehdyssä kirjauksessa.

Aineistolähtöinen sisällönanalyysi voidaan toteuttaa temaattisella analyysillä, jossa koodataan aineistosta tunnistetut teemat. Koodaamisella tarkoitetaan laajemman tekstin yksinkertaistamista tai uudelleen luokittelua, aineiston osien erottelua ja uudelleen yhdistämistä yhteisen ominaisuuksien perusteella taikka pelkistettyjen ilmaisujen tunnistamista, kuten sanoja, sanapareja, kuvauksia (Juhila, n.d.a). Temaattista analyysiä voidaan toteuttaa joko *induktiivisesti* eli aineistolähtöisesti tai *deduktiivisesti* eli teorialähtöisesti (Juhila, n.d.b)

Temaattinen analyysi on hyvin joustava laadullisen tutkimuksen menetelmä, joka soveltuu datassa esiintyvien toistuvien teemojen tai tapahtumien tunnistamiseen, analysointiin sekä raportointiin. Joustavuutensa ja ennalta määritettyyn teoriaan sitoutumattomana menetelmänä temaattinen analyysi soveltuu hyvin myös kokemusten ja asioiden merkityksen tutkimiseen. Temaattinen analyysi voidaan toteuttaa joko aineistolähtöisesti eli induktiivisesti taikka teorialähtöisesti eli deduktiivisesti. (Braun & Clarke, 2006, s. 12; Nowell ym., 2017, s. 2)

Temaattisen analyysin vaiheita ovat: aineistoon tutustuminen, alustavien koodien muodostaminen, teemojen etsiminen, teemojen tarkastelu, teemojen yksilöinti ja nimeäminen sekä raportointi. Ennen temaattisen analyysin toteutusta on päätettävä, onko analyysi aineisto- vai teorialähtöinen ja toteutettavan koodauksen tapa eli koodataanko havaintojen yksityiskohtia (kuvaileva koodaus) vai koodataanko teeman näkökulmia (tulkitseva koodaus). Koodaamisen osalta on ennakkoon päätettävä myös siitä, että koodataanko täsmällistä merkitystä vai piilomerkitystä. Ennakolta on myös päätettävä, että mitkä kaikki asia tulkitaan teemoiksi ja millainen merkitys teeman esiintymistajuudelle annetaan tulosten tulkinnassa. (Braun & Clarke, 2006, ss. 9–11; Elomaa-Krapu, n.d.)

Raportoinnin ja tilannekuvan kehittämiseen liittyvä tutkimusosuus toteutettiin laadullisena tapaustutkimuksena ja työssä hyödynnettiin viranomaisen valmista aineistoa eli järjestelmään tehtyjä kirjauksia. Asiakasyhteydenottojen sisällön analysointi toteutettiin teorialähtöisenä

temaattisena analyysinä, jossa seurattavat teemat olivat lainsäädännön kiellettyjä syrjintäperusteita. Analyysivaiheessa yhteydenottojen kuvaustekstistä poimittiin kaikkia syrjintäperusteita kuvaavat sanat tai ilmaisut. Teemojen koodaus toteutettiin yksityiskohtia korostaen ja täsmällisen merkityksen tasolla, jotta koodausta voitiin myöhemmässä vaiheessa hyödyntää tilannekuvan muodostamisessa.

4.3 Puolistrukturoitu ryhmähaastattelu

Kehittämistoiminnan tiedonkeruumenetelmiin kuuluvat erilaiset haastattelut, kuten lomaketutkimuksen kaltainen strukturoitu haastattelu taikka avoin haastattelu, johon ei välttämättä laadita ennakkoon täsmällisiä kysymyksiä. Näiden kahden välimaastoon sijoittuu teemahaastattelu, jossa haastateltavien kanssa keskustellaan ennalta päätetyistä aiheista, mutta aiheiden järjestys sekä keskusteluun käytetty aika vaihtelee haastateltavittain (Eskola ym., 2018, ss. 28–30)

Haastattelut voidaan toteuttaa siten, että jokaisen haastateltavan kanssa keskustellaan erikseen tai haastattelu voidaan toteuttaa yhdellä kertaa ryhmänä (Hyvärinen ym., n.d.). Ryhmähaastattelu soveltuu tilanteisiin, joissa tavoitellaan haastateltavien vastauksista yhdessä tuotettua kollektiivista käsitystä asiasta, joka ei kosketa haastateltavan henkilökohtaisia asioita (Saaranen-Kauppinen ym., 2009, ss. 57–59).

5 Suunnittelu ja toteutus

Raportointityökalun kehitystyössä sovellettiin CRISP-tiedonlouhintaprosessia (kuva 5), jonka ensimmäiseen vaiheeseen liittyvät kappaleet 5.1 Raportointityökalun määrittely ja 5.2 Asiakasyhteydenottojen sisällönanalyysin suunnittelu. CRISP:n vaiheisiin 2–4 liittyy kappale 5.3 Raportointityökalun toteutus. Kehitystyö päättyi siihen, kun uusi raportointityökalu ja sen käyttöä koskeva ohjeen esiteltiin toimeksiantajalle.

5.1 Raportointityökalun määrittely

Kehitystyön ennakkoselvitysvaiheessa alkanut keskustelu toimeksiantajan tarpeista ja työn tavoitteista viimeisteltiin toteuttamalla toimeksiantajan asiantuntijoille puolistrukturoitu ryhmähaastattelu. Ryhmähaastatteluun osallistui kaksi asiantuntijaa ja keskustelu litteroitiin. Ryhmähaastattelussa keskityttiin toimeksiantajan liiketoimintamallin ymmärtämiseen eli asiakasyhteydenottojen analysoinnin tarpeisiin ja syvyyteen. Haastattelun perusteella

täsmennettiin kehittämistyön tavoitteita sekä priorisoitiin niiden toteutusjärjestystä. Tavoitteet on käsitelty aiemmin kappaleessa 3.

Ryhmähaastattelun perusteella täsmentyi myös raportointityökalun toteutustaso. Haastattelun aikana esiin tulleet raportointityökalun määrittelyt on kirjattu soveltuvin osin taulukoon 2 (Ailimpieti & Ruottinen, 2002a, 2002b). Toimeksiantaja halusi, että uusi raportointityökalu tehtäisiin tässä kohtaa MVP-tuotteena (*Minimum Viable Product*, suomeksi pienin julkaistava tuote). Toimeksiantajalle tärkeitä ominaisuuksia ovat, että raportointityökalu on toiminnallisuudeltaan nopeasti käytettävissä ja tilannekuva nopeasti päivitettävissä. Toimeksiantaja halusi, että tilannekuvassa esitetään visuaalisesti rakenteellinen data. Toimeksiantaja tarvitsi myös ohjeen raportointityökalun käyttöön.

Taulukko 2. MVP-tuotteen ominaisuuksia ryhmähaastattelun perusteella.

<i>Määrittelyt</i>	
<i>MVP:n kuvaus</i>	Raportointityökalu, jonka käyttö vaatii vähän toimenpiteitä ja joka tuottaa visuaalisen kokonaiskuvan asiakasyhteydenotoista.
<i>Asiakkaat</i>	Työsuojeluviranomaisten raportointitehtävissä toimivat henkilöt, esihenkilöt ja johtajat.
<i>Mitä ongelmaa ratkaistaan ja millainen asiakkaan polku on kehitystyön aikana</i>	Asiakasyhteydenottojen määrän raportointi on liian hidasta ja sitoo tarpeettoman paljon henkilötyöpäiviä. Syrjintäperusteita koskevaa tietoa ei saada seurattua. Tarvittava data ei ole valvontatietojärjestelmästä rakenteisena datana, joten raportoinnin yhteydessä tieto etsitään lukemalla asiakastietoja.
<i>Tarvittavat toiminnallisuudet, ominaisuudet ja ulkoiset vaatimukset</i>	Visuaalisen tilannekuvan tulee olla nopeasti päivitettävissä. Visuaalisen tilannekuvan tukena on oltava ristiintaulukoitua numeerista tietoa. Raportoinnissa eikä tilannekuvassa käsitellä henkilötietoja.
<i>Odotettu lopputulos</i>	Raportointityökalu Excel-tiedostona, jossa rakenteinen data kuvataan visuaalisena tilannekuvana. Asiakasyhteydenoton kuvauksesta muodostettu rakenteista dataa. Loppukäyttäjälle ohje raportointityökalun käyttöön.
<i>Hinta ja aikataulutus</i>	Aikataulutus: käytettävissä vuoden 2023 vuosiraportointiin eli tammi-maaliskuun 2024 aikana.

Ryhmähaastattelussa ilmeni, että toimeksiantajalle riittää se, että raportoinnissa ja tilannekuvassa keskitytään kuvailevaan analytiikkaan. Raportointityökalun käyttötarkoitus on helpottaa ja nopeuttaa raportoinnin toteutumista sekä seurata ilmiön kokonaiskehitystä. Ryhmähaastattelussa määritettiin kehitystyön toteutusaikataulu. Aikataulusyistä toimeksiantaja teki päätöksen, että raportointityökalun ensimmäinen MVP- versio toteutetaan

Excel-ratkaisuna. Tämä ratkaisu mahdollisti uuden raportointityökalun käyttöönoton maaliskuussa 2024 aikana.

5.2 Asiakasyhteydenottojen sisällönanalyysin suunnittelu

Asiakasyhteydenottojen kuvaus on rakenteetonta dataa eli vapaata tekstiä, jolla ei ole ennalta määritettyä vakioitua tai standardoitua sisältöä tai tiedon esitysjärjestystä.

Asiakasyhteydenottojen kuvauksista tarkasteltiin datan laatutekijöistä erityisesti sitä, että soveltuuko kuvausdata analyysityöhön.

Sisällönanalyysi sisälsi useita lukukierroksia, joiden aikana toteutetaan teorialähtöinen temaattinen analyysi. Kuvauksista tutkittiin kiellettyjä syrjäntäperusteiden esiintymistä ja kuvaustekstistä poimittiin jokaista syrjäntäperustetta kuvaavat sanat syrjäntäperusteittain. Tutkimuksessa hyödynnettiin taulukon 3 mukaista taulukointia. Taulukossa 3 on esitetty vain neljä ensimmäistä syrjäntäperustetta esimerkkinä.

Taulukko 3. Esimerkki tutkimuksessa hyödynnettävästä ristiintaulukoinnista.

Tutkimusta varten luotu juokseva numero	Kuvausteksti	Alkuperä	Ihönväri	Ikä	Kansalaisuus
1	Lorem ipsum dolor sit amet, consectetur adipiscing elit...	Syrjäntäperustetta kuvaavat sanat tai tieto, ettei niitä tunnistettu	Syrjäntäperustetta kuvaavat sanat tai tieto, ettei niitä tunnistettu	Syrjäntäperustetta kuvaavat sanat tai tieto, ettei niitä tunnistettu	Syrjäntäperustetta kuvaavat sanat tai tieto, ettei niitä tunnistettu

Suunnitteluvaiheessa tunnistettiin kaksi vaihtoehtoista tapaa toteuttaa sisällönanalyysi. Ensimmäinen vaihtoehto oli toteuttaa analyysi ilman analyysiohjelmistoa eli käytännössä aineistoa lukemalla ja manuaalisesti merkintöjä tekemällä. Toinen vaihtoehto oli hyödyntää mm. Excelin tai Power BI:n tekoälypohjaisia tekstianalyysitoimintoja tai muita kaupallisia tekstianalyysiohjelmia sillä edellytyksellä, että tekoälyn hyödyntäminen on ylipäänsä mahdollista tutkimusaineiston salassapidon takia.

5.3 Raportointityökalun toteutus

Raportointityökalun kehitystyö toteutettiin suunnitellusti kolmessa eri vaiheessa:

1. Rakenteellisen datan analysointi ja visuaalinen tilannekuva
2. Asiakasyhteydenottojen sisällönanalyysi
3. Rakenteen lisääminen asiakasyhteydenottojen tekstidataan

Raportoinnin kehitystyö sekä yhteydenottojen sisällön analysointi toteutettiin CRISP-tiedonlouhintaprosessin mukaisesti. Liiketoimintatarpeiden ymmärtämistä koskeva vaihe toteutettiin tämän opinnäytetyön suunnitteluvaiheen aikana. Tämän vaiheen havainnot on esitetty edellä kappaleessa 5.1.

Kehitystyön ensimmäisessä vaiheessa katselmoitiin datan laatutekijät ja erityisesti sitä, että soveltuuko data tarkoitukseensa (Väre, 2019, s.198). Tällä toteutettiin CRISP:n datan ymmärtämistä koskeva vaihe.

Toisessa vaiheessa toteutettu asiakasyhteydenottojen sisällönanalyysi toteutti CRISP:n datan esikäsittelyvaihetta. Tämän vaiheen aikana tunnistettiin sanoja tai sanapareja, joilla henkilön kokema syrjintä oli kirjattu valvontatietojärjestelmään. Toisen vaiheen aikana tutkittiin asiakasyhteydenotoista myös toimeksiantajan muita nimeämiä kokonaisuuksia.

Kolmannessa vaiheessa asiakasyhteydenottojen dataan pyrittiin lisäämään rakennetta ja tällä tavalla toteutettiin CRISP:n mallinnusvaihetta. Mallinnuksessa käytetyissä Excel-funktioissa hyödynnettiin toisen vaiheen aikana tunnistettuja sanoja tai sanapareja. Mallinnusvaiheessa myös testattiin käytettyjä hakusanoja ja seurattiin väärin positiivisten havaintojen syntymistä. Väärin positiivisten havaintojen selvittäminen liittyy hakusanan validiteettiin eli siihen, että se kuvaa sitä, mitä oli tarkoituskin kuvata (Tilastokeskus, n.d.).

5.3.1 Datan laadun arviointi

Ensimmäisen vaiheen aluksi tarkasteltiin tutkimusaineistosta datan laatutekijöitä kehittämistyön tavoitteiden näkökulmasta. Datan laadun osatekijät ovat: saatavuus, analysoitavuus, ajantasaisuus, soveltuvuus käyttöönsä, tarkkuus, täydellisyys sekä paikkansapitävyys. Datan laatua arvioitiin siten, että laatu on joko riittävällä tasolla tai siinä on puutteita (taulukko 4).

Taulukosta 4 nähdään, että tutkimusaineiston datan laatu vaihtelee osatekijöittäin. Data on saatavissa ja analysoitavissa eli data on siirrettävissä valvontatietojärjestelmän ulkopuolelle muodossa, joka mahdollistaa analysoinnin. Data on lataushetkellä täysin ajantasaista ja paikkansa pitävää.

Visuaalinen tilannekuva hyödyntää järjestelmässä olevaa rakenteellista dataa, jolloin datan tulee olla myös tarkkaa ja täydellistä. Tarkkuuden ja täydellisyyden osalta datan laadusta todettiin puutteita. Tutkimusaineistosta puuttui kokonaan toimialatieto 37 yhteydenotosta (7 %), kuvaus viidestä (1 %) ja työnantajan koko 129 yhteydenotosta (25 %).

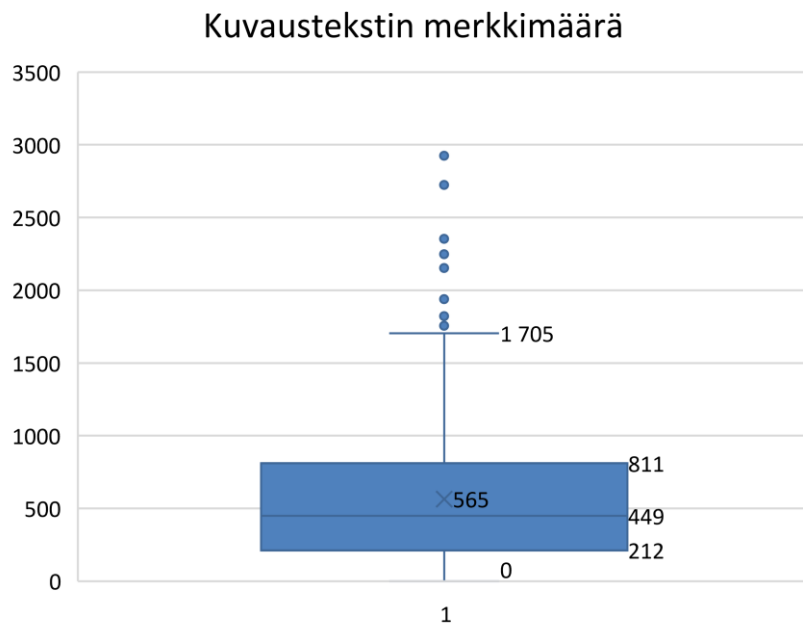
Toimeksiantajan ryhmähaastattelussa varauduttiin myös siihen, että datan laadussa havaitaan puutteita. Haastattelun aikana keskusteltiin siitä, miten suhtaudutaan puutteellisiin ja epätarkkoihin tietoihin. Toimeksiantajan kanta oli, että puutteellisia tai epätarkkoja yhteydenottoja ei poisteta tutkimusaineistosta vaan tyhjä arvo korvataan tiedolla ”tieto puuttuu”. Tällä ratkaisulla turvattiin se, että saadaan luotua mahdollisimman tarkka kuva yhteydenottojen syrjintäperusteista.

Taulukko 4. Kehitystyössä tarvittavan datan laadun arviointi.

<i>Datan laadun osatekijä</i>	<i>Arvioinnin tulos</i>
<i>Saatavuus</i>	Riittävä
<i>Analysoitavuus</i>	Riittävä
<i>Ajantasaisuus</i>	Riittävä
<i>Soveltuvuus käyttöön</i>	Puutteellista
<i>Tarkkuus</i>	Puutteellista
<i>Täydellisyys</i>	Puutteellista
<i>Paikkansa pitävyys</i>	Riittävä

Rakenteettoman datan soveltuvuus käyttöön analysoinnin näkökulmasta todettiin puutteelliseksi, vaikka ne ovat valvontatyön toteuttamisen kannalta täysin tarpeellisia ja valvontatyöhön soveltuvia. Kuvaustekstit olivat hyvin vaihtelevia sisällöllisesti ja pituudeltaan. Viidestä yhteydenotosta kuvausteksti puuttui kokonaan. Näiden jälkeen kuvaustekstien pienin merkkimäärä oli 5 ja suurin merkkimäärä 2924, aineiston kuvausten merkkimäärän keskiarvo oli 565. Kuvasta 6 nähdään myös, että puolessa yhteydenotoista kuvaustekstin merkkimäärä oli 212 ja 811 merkin väliltä. Vertailukohteeksi, tämän kappaleen merkkimäärä tämä virke mukaan luettuna on 706 merkkiä, kun välimerkit ja väliluonnit lasketaan mukaan.

Kuva 6. Yhteydenoton kuvaustekstin merkkimäärän vaihtelu.



Yhteydenoton kuvauksista havaittiin, ettei kirjaus aina sisältänyt tietoa tarkemmasta syrjintäperusteesta. Sen sijaan oli kirjattu pelkästään tarkastajan muistiinpanoiksi luokiteltavaa tietoa. Esimerkiksi vain tieto siitä, että työntekijä on tehnyt valvontapyynnön tai tieto työntekijän yhteydenotosta, jossa työntekijä halusi neuvoja esimerkiksi turvasähköpostin lähettämisessä.

5.3.2 Visuaalinen tilannekuva

Ensimmäisen vaiheen aikana toimeksiantajalle toteutettiin visuaalinen tilannekuva niistä tiedoista, jotka olivat valmiiksi rakenteellisena datana. Esimerkiksi kaavio vastaanotettujen asiakasyhteydenottojen määrästä kuukausittain sekä kaavio yhteydenottoon merkittyjen asialuokkien määrästä. Ensimmäinen versio tilannekuvasta toimitettiin toimeksiantajalle jo tammikuussa 2024, jonka jälkeen toimeksiantaja testasi raporttipohjaa. Raportointityökalua muokattiin saadun palautteen jälkeen. Lopullinen raportointityökalu ja siihen liittyvä ohje toimitettiin toimeksiantajalle maaliskuussa 2024.

Toimeksiantajan ratkaisun mukaisesti visuaalinen tilannekuva toteutettiin Excel-työkirjana. Toimeksiantaja ei tarvitse reaaliaikaista tilannekuvaa, joten toteutusmalliksi valikoitiin ratkaisu, jossa toimeksiantaja lataa tiedot valvontatietojärjestelmästä ja lisää ne manuaalisesti raportointityökaluun. Datan lisäämisen jälkeen raportointityökalussa olevat taulukot ja kaaviot päivittyvät.

Kehitystyön aikana tunnistettiin mahdollisuus laajentaa tilannekuva koskemaan myös muita yhteydenotoista ilmeneviä asialuokkia, joten tilannekuvaan toteutettiin toiminnallisuutena mahdollisuus tietojen osittaiseen esittämiseen asialuokakohtaisesti kokonaiskuvan lisäksi. Raportointityökalun ja sitä kautta tilannekuvaan toteutettu ratkaisu luo samanaikaisesti näkymän neljän asialuokan yhteydenottoihin. Nämä asialuokat ovat: työsyrintä, häirintä ja epäasiallinen kohtelu, psykososiaalinen kuormitus sekä seksuaalinen häirintä.

5.3.3 Asiakasyhteydenottojen sisällönanalyysi

Asiakasyhteydenottojen sisällönanalysointi toteutettiin toimeksiantajan laitteilla ja tietotekniikkaympäristössä analysoitavan aineiston salassapitovaatimuksien takia. Sisällönanalysointi päätettiin toteuttaa ilman kaupallisten analyysiohjelmistojen käyttöä. Asiakasyhteydenottojen sisällönanalyysi toteutettiin täysin manuaalisena työnä ilman kaupallisia analyysiohjelmistoja sekä ilman Excelin tekoälypohjaisten apuohjelmien hyödyntämistä.

Ratkaisu sisällönanalyysin manuaalisesta toteutuksesta perustui pitkälti Aluehallintovirastojen joulukuussa 2023 julkaisema tekoälyn käyttöä koskevaan linjaukseen. Linjaus on, että tekoälyä saa käyttää vain julkisten asioiden käsittelyyn. Linjauksessa korostettiin myös sitä, ettei tekoällylle saa antaa henkilötietoja. (Aluehallintovirasto, 2023). Asiakasyhteydenotot ovat salassa pidettävä tietoa lain työsuojelun valvonnasta ja työpaikan työsuojelutoiminnasta (44/2006) 10 §:n perusteella. Tämän lisäksi asiakasyhteydenotot sisältävät muita salassa pidettäviä tietoja kuten terveydentilatietoja. Tekoälylinjauksen lisäksi ratkaisuun vaikutti toimeksiantajan käytettävissä olevat tietojärjestelmät sekä tilanne, ettei tietoja saa siirtää EU/ETA-alueen ulkopuolelle. Manuaalista sisällönanalysointityötä puolsi myös tarve selvittää tutkittavan datan laatua ja erityisesti käyttötarkoitukseen soveltumista.

Sisällönanalyysin ensimmäisellä lukukierroksella keskityttiin arvioimaan datan laatua kuvaustekstin osalta. Pääosa kirjauksista oli tehty näkökulmasta, jossa korostuu tarkastajan rooli tiedon kirjaajana. Teksteissä toistuu kirjaus ”asiakas soittaa” tai ”työntekijä kertoo”. Osaan kuvauksia oli tehty vain sellainen kirjaus, jossa kerrottiin tarkastajan tekemiä toimenpiteitä, tai millaisia ohjeita tarkastaja oli asiakkaalle antanut taikka kuvaukseen oli ytimekkäästi kirjattu vain ”työsyrintä” ilman lisätietoja. Kuvaustekstien soveltuvuus analyysityöhön todettiin puutteelliseksi ja sen takia datan laatu on puutteellista.

Toinen lukukierros vahvisti käsitystä datan laadun puutteista. Toisella lukukierroksella keskityttiin tunnistamaan kuvaustekstistä syrjinnän esiintymismuotoja kuvaavia sanoja tai ilmaisuja. Oli mahdollista, että yhdestä kirjauksesta ilmenisi useita syrjintäperusteita ja

kuvaustekstistä tunnistettiin kaikki syrjintäperusteet. Syrjintäperusteiden koodaus toteutettiin yksityiskohtia korostaen ja täsmällisen merkityksen tasolla, jotta koodausta voitaisiin myöhemmässä vaiheessa hyödyntää tilannekuvan muodostamisessa.

Tutkimusaineisto koostui 512 asiakasyhteydenoton tiedoista, jotka työsuojeluviranomainen oli vastaanottanut vuonna 2023. Taulukosta 5 nähdään, että 30 %:ssa kirjauksia kuvaustekstistä ei pystytty tunnistamaan yhtään syrjintäperustetta. Noin 50 %:ssa kirjauksia ilmeni vain yksi syrjintäperuste. Moniongelmaisina tapauksina voidaan pitää niitä, joissa ilmenee useita syrjintäperusteita samanaikaisesti. Tällaisia aineistosta 3–5 eri syrjintäperusteen tapauksia aineistosta löytyi yhteensä 20 kpl.

Taulukko 5. Yhteydenotosta ilmenevien syrjintäperusteiden määrä.

<i>Kuvauksesta ilmenevien syrjintäperusteiden määrä</i>	<i>Määrä</i>	<i>Osuus</i>
<i>1 syrjintäperuste</i>	253	49,4 %
<i>2 syrjintäperustetta</i>	83	16,2 %
<i>3 syrjintäperustetta</i>	12	2,3 %
<i>4 syrjintäperustetta</i>	7	1,4 %
<i>5 syrjintäperustetta</i>	1	0,2 %
<i>Ei yhtään syrjintäperustetta</i>	156	30,5 %
<i>Yhteensä</i>	512	100,0 %

Asiakasyhteydenottojen kirjauksista kerättiin kappaleessa 5.2 kuvatun taulukon mukaisesti jokaista syrjintäperustetta kuvaavat sanat siinä muodossa kuin ne kuvaustekstissä esiintyi. Ennako-odotusten vastaisesti kirjauksista ei ilmennyt tapa, jolla syrjintä esiintyi vaan sen sijaan tilanneyhteys. Aineisto sisälsi vain yksittäisiä kirjauksia syrjinnän tavasta, joten tämän opinnäytetyön tuloksia ei pystytä reflektoimaan oikeusministeriön vihapuhetta koskevan tutkimuksen tulosten kanssa. Asiakasyhteydenottojen kirjauksissa korostuivat tilanteet, joissa syrjintää oli koettu tai tilanne, johon syrjinnän kokemus liittyi. Esimerkiksi terveydentilaan liittyvän syrjintäperusteen kohdalla toistuivat muun muassa sanat sairausloma, sairauslomalla ja terveydentilan.

Taulukossa 6 esitetään kaikkien syrjintäperusteiden osalta analyysityössä tunnistetut tapausmäärät. Tuloksia tarkastellessa on muistettava, että yksi yhteydenotto on voinut sisältää viitteitä useista syrjintäperusteista. Tulokset esitetään syrjintäperusteittain. Taulukko 6 sisältää myös esimerkkejä siitä, millaiset sanat on yhdistetty kyseiseen

syrijintäperusteeseen. Esimerkkeiksi on poimittu sellaisia, jotka ovat toistuneet useamman kerran.

Asiakasyhteydenotoista tunnistettiin eniten terveydentilan liittyviä yhteydenottoja, joita löytyi yhteensä 147. Toiseksi eniten löytyi tapauksia, joissa syrjäntäperuste oli muu henkilöön liittyvän syy (101) ja kolmanneksi eniten alkuperään liittyviä yhteydenottoja (54).

Mielipiteeseen, perhesuhteisiin, poliittiseen toimintaan, seksuaaliseen suuntautumiseen, uskuntoon tai vakaumukseen liittyviä yhteydenottoja tunnistettiin yksittäisiä.

Taulukko 6. Asiakasyhteydenotoista ilmenevien syrjäntäperusteiden tapausmäärä

<i>Syrjäntäperuste</i>	<i>Tapausmäärä</i>	<i>Esimerkkejä poiminnoista</i>
<i>Alkuperä</i>	54	alkuperä, alkuperänsä, maahanmuuttaja, ulkomaalainen, rassistisesti, rassistista
<i>Ammattiyhdistystoiminta</i>	11	ammattiyhdistystoimintansa, luottamusmiehenä
<i>Ikä</i>	26	ikä, iän, ikäsyrjäntä, nuori, eläkkeellä
<i>Kansalaisuus</i>	42	kansalaisuuden, ukrainalainen, filippiiniläisistä
<i>Kieli</i>	19	kieli, kielen, vain suomea,
<i>Mielipide</i>	alle 10	mielipide, mielipiteensä
<i>Muu henkilöön liittyvä syy</i>	101	muun henkilöön liittyvä, tuonut esiin epäkohtia
<i>Perhesuhteet</i>	alle 10	vaimo, mies, puoliso, perhe
<i>Poliittinen toiminta</i>	alle 10	poliittisesti, poliittisen
<i>Seksuaalinen suuntautuminen</i>	alle 10	seksuaalisesta suuntautumisesta
<i>Sukupuoli</i>	44	sukupuolen, perhevapaa, raskaus, vanhempainvapaa
<i>Terveydentila</i>	147	terveydentila, sairausloma, sairauslomalla, sairaudestaan
<i>Uskonto</i>	alle 10	uskonto, rukoilu
<i>Vakaumus</i>	alle 10	vakaumus, vakaumuksensa
<i>Vammaisuus</i>	15	vammansa, näkövammaisuus, kuulovamma

Sisällönanalyysin aikana tutkittiin myös sitä, että tapahtuiko syrjintä työnantajan tai hänen edustajansa eli esihenkilön toimesta vai toisen työntekijän tai työyhteisön toimesta. Tämä on uusi näkökulma raportoinnissa, eikä sitä ole aiemmin tällä tasolla työsuojelussa seurattu eikä raportoitu. Tutkimusaineistosta löytyi yhdeksän yhteydenottoa, joista oli tulkittavissa tilanne, että työyhteisöön kuuluvat työkaverit syrjivät henkilöä, joka oli ollut yhteydessä työsuojeluviranomaiseen. Pääasiallisesti syrjintä koettiin työnantajan tai työnantajan työjohto-oikeutta käyttävän esihenkilön suunnasta.

5.3.4 Rakenteisuuden lisääminen asiakasyhteydenottoja koskevaan dataan

Kuten edellä todettiin, asiakasyhteydenottojen kuvausteksteistä tunnistettiin syrjinnän eri esiintymismuotoja ja jokaisesta kuvauksesta kirjattiin kaikki syrjintäperusteita kuvaavat sanat. Analyysityön aikana poimituista sanoista luotiin hakusanoja, jotka olivat käytännössä merkkijonoja (string) eli aina kyseessä ei ollut kieliopillisesti kokonainen sana vaan sen osa. Esimerkiksi terveydentilaan liittyvän syrjintäperusteen kohdalla esiintyi muun muassa sanat sairausloma, sairauslomalla, sairauslomalta, sairauslomia jne. Kuvausteksteissä yksi sana pystyi esiintymään useita kertoja samassa tai eri sijamuodossa.

Poimituista sanoista tunnistettiin niille yhteisiä muotoja eli sanoja typistettiin. Esimerkiksi terveydentilaan liittyviä sanat *sairausloma*, *sairauslomalla*, *sairauslomalta*, *sairauslomia* ovat yhteiseltä juurimuodoltaan *sairauslom*. (Harjunpää, 2004, ss. 5–6) Juurimuodot tunnistettiin manuaalisesti. Juurimuotoja hyödynnettiin Excel-funktiossa, joilla tunnistettiin syrjintäperusteen hakusanan esiintyminen kuvaustekstissä.

Luodut hakusanat testattiin väärin positiivisten havaintojen tunnistamiseksi. Väärät positiiviset havainnot tunnistettiin vertaamalla Excel-funktion tuloksia manuaalisesti tehdyn sisällönanalyysin luokitteluun. Väärä positiivinen havainto tarkoittaa tilannetta, jossa hakusana aktivoi syrjintäperusteen luokitteluosuman, vaikkei yhteydenotto koske kyseistä syrjintäperustetta. Esimerkiksi yhteydenotot ammattiyhdistystoimintaan liittyvästä syrjinnästä sisälsi usein sanan luottamusmies ja tästä muodostui hakusana kyseiselle syrjintäperusteelle. Tällä hakusanalla syntyi väärä positiivisia havaintoja ammattiyhdistystoiminnalle tilanteissa, joissa yhteydenoton kuvaukseen oli kirjattu esimerkiksi, että työpaikalla ole luottamusmiestä tai luottamusmies oli kysynyt neuvoa työpailla esiintyneeseen tilanteeseen. Samalla tavalla ikään liittyvä syrjintäperuste saa väärä positiivisia havaintoja hakusanalla ikä, kun kirjaus sisältää sanoja mikä, mikään tai eikä.

Väärä positiivisia havaintoja esiintyi lähes jokaisen syrjintäperusteiden kohdalla. Jos väärä positiivisia havaintoja ei syntynyt niin kyseiseen syrjintäperusteeseen liittyi vain yksittäisiä

yhteydenottoja. Jos syrjintäperusteeseen liittyi vain muutamia kuvauksia niin syrjintäperustetta kuvaavia hakusanoja ei muodostunut yhtä tai kahta kappaletta enempää. Yksittäiset hakusanat eivät riitä kuitenkaan kuvaamaan läheskään tyhjentävästi syrjintäperustetta.

Hakusanojen testauksen osoitti, että syrjintäperusteiden hakusanoilla esiintyi vääriä positiivisia havaintoja. Väärät positiiviset havainnot johtivat tilanteisiin, joissa yhteydenotto luokiteltiin väärään syrjintäperusteeseen liittyväksi. Tutkimuksen lopputulos on, ettei kuvaustekstiin saatu lisättyä tavoitteena ollutta lisärakennetta, koska sisällönanalyysin ei tuottanut luotettavia hakusanoja syrjintäperusteiden luokitteluun.

5.4 Kehitystyö tulosten yhteenveto

Tämän opinnäytetyön kehitysosuuden aikana muodostui kokonaiskuva siitä, millaisista työsyryntätilanteista oltiin yhteydessä työsuojeluviranomaiseen vuonna 2023.

Toimeksiantajalle kehitettiin raportointityökalu tilannekuvan muodostamista varten sekä laadittiin ohje raportointityökalun käyttöä varten. Valmistunut raportointityökalu täyttää sille asetetut tavoitteet suurelta osin (taulukko 7). Valmistunut raportointityökalu mahdollistaa kaiken kaikkiaan neljän eri asialuokan asiakasyhteydenottojen tilannekuvan muodostamisen, mikä on enemmän kuin alkuperäinen tavoite. Kehitystyön aikana muodostui myös kokonaiskäsitys toimeksiantajan käytettävissä olevan datan laadusta.

Työssä tutkittiin vuonna 2023 työsuojeluviranomaiselle tehtyjä asiakasyhteydenottoja, jotka liittyivät työsyryntään. Kehitystyön aikana toteutettu sisällönanalyysi tuotti kokonaiskuvan vuonna 2023 vastaanotettujen yhteydenottojen syrjintäperusteista. Tämä ei ollut työn nimenomaisena tavoitteena. Tietämys siitä, millaisia syrjintäperusteita muodostui vuoden 2023 asiakasyhteydenotoista on syntynyt manuaalisesti tehdyn sisällönanalysoinnin tuloksena kehitystyön sivutuotteena.

Tutkimusaineisto sisälsi yhteensä 512 asiakasyhteydenoton tiedot. Kolme tyypillisintä tunnistettua syrjintäperustetta olivat terveydentila (147), muu henkilöön liittyvä syy (96) ja alkuperä (54). Vähiten löytyi yhteydenottoja seuraavista syrjintäperusteista: mielipide, perhesuhteet, poliittinen toiminta, seksuaalinen suuntautuminen, uskonto ja vakaumus.

Merkittävä havainto on se, että osasta tutkimusaineiston yhteydenottoja ei tunnistettu yhtään syrjintäperustetta. Näiden tapausten osuus oli 30 %. Tällöin oli mahdollisesti tehty kirjaus, jossa kerrottiin tarkastajan tekemiä toimenpiteitä, tai millaisia ohjeita tarkastaja oli asiakkaalle antanut taikka kuvaukseen oli ytimekkäästi kirjattu vain ”työsyryntä” ilman lisätietoja.

Tyypillisesti yhdestä yhteydenotosta tunnistettiin yksi syrjäntäperuste, mutta yhteydenotto saattoi sisältää viitteitä jopa viiden eri syrjäntäperusteen samanaikaisesta olemassaolosta. Puolet yhteydenotoista sisälsi yhden syrjäntäperusteen ja moniongelmaisia eli 3–5 syrjäntäperusteen yhteydenottoja oli koko aineistossa yhteensä 20 (taulukko 5).

Tässä kohtaa on tarpeellista korostaa sitä, että yhteydenottojen kuvaukset antavat vain *viitteitä* työsyrynnän esiintymiselle, koska kaikki yhteydenotot eivät johda tarkastukseen. Tutkimusaineisto sisälsi syrjäntäolettan täyttäviä ja valvontatoimenpiteisiin johtaneita ilmoituksia sekä sellaisia yhteydenottoja, joiden perusteella työsuojeluviranomainen ei ollut ryhtynyt valvontatoimenpiteisiin. Merkittävään osaan tutkimusaineiston yhteydenottoja oli merkitty, ettei asia johda jatkotoimenpiteisiin. Tällöin yhteydenotto on johtanut asiakkaan neuvontaan tai tilanne on merkitty tiedoksi. Niistä tutkimusaineiston yhteydenotoista, joihin oli merkitty jatkotoimenpiteisiin johtaminen, ei ole tietoa tosiasiallisten jatkotoimenpiteiden toteuttamisesta.

Työn tavoitteena oli tuottaa toimeksiantajalle raportointityökalu, jolla toimeksiantaja pystyy luomaan työsyrynnän asiakasyhteydenotoista ajantasaisen tilannekuvan nopeasti ja vähin toimenpitein. Tämän toteuttamiseksi tavoitteena oli luoda lisärakennetta asiakasyhteydenottojen kuvausdataan, jotka on tallennettu vapaana tekstinä tietojärjestelmään. Asiakasyhteydenottojen kuvausteksteihin toteutettiin suunnitellusti teorialähtöinen temaattinen sisällönanalyysi, jonka tuloksia hyödyntämällä oli tarkoitus lisätä rakenteellisuutta asiakasyhteydenoton kuvausdataan.

Yhteydenottojen kuvaustekstejä analysoitaessa tunnistettiin syrjäntäperusteita koskevia hakusanoja, mutta testausvaiheessa niiden havaittiin aktivoivan runsaasti vääriä positiivisia havaintoja. Väärät positiiviset havainnot aktivoivat väärän syrjäntäperusteen, jolloin tiedonluokittelu tapahtui väärin. Tutkimuksen aikana ei siten tunnistettu luotettavia hakusanoja syrjäntäperusteille. Tämän havainnon perusteella päätettiin, että tilannekuvassa hyödynnetään vain valmiiksi rakenteellisena olevaa dataa. Päätös johti myös siihen, ettei työsyryntää koskevista asiakasaloitteista saada luotua täysimääräisesti tavoitteena ollut kokonaiskuvaa.

Raportointityökalun kehittämislle asetetut tavoitteet saavutettiin osittain (taulukko 7). Raportointityökalu on toteutettu toimeksiantajan määritysten mukaisesti Excel-työkirjana. Raportointityökalu sisältää myös visuaalisen tilannekuvan yhteydenotoista. Tilannekuvassa on hyödynnetty vain valmiiksi rakenteellisena olevaa dataa, joista on tehty ristiintaulukoita numeerisen tiedon esittämiseksi sekä kaavioita visualisoimaan asioiden kehitystä toimeksiantajan määritysten mukaisesti. Tilannekuva sekä sen ristiintaulukot ja kaaviot

päivittyvät helposti ja nopeasti sen jälkeen, kun toimeksiantaja on lisännyt yhteydenotot raportointityökaluun. Nykyisen raportointityökalun ja tilannekuvan sisältö on muodostettu hyvin pitkälti vakiintuneen vuosiraportoinnin sisällön perusteella.

Taulukko 7. MVP-tuotteen määrittelyjen toteutumisen arviointi.

	<i>Määrittelyt</i>	<i>Määrittelyjen toteutuminen</i>
<i>MVP:n kuvaus</i>	Raportointityökalu, jonka käyttö vaatii vähän toimenpiteitä ja joka tuottaa visuaalisen kokonaiskuvan asiakasyhteydenotoista.	Toteutui osittain
<i>Asiakkaat</i>	Työsuojeluviranomaisten raportointitehtävissä toimivat henkilöt, esihenkilöt ja johtajat.	
<i>Mitä ongelmaa ratkaistaan ja millainen asiakkaan polku on kehitystyön aikana</i>	Asiakasyhteydenottojen määrän raportointi on liian hidasta ja sitoo tarpeettoman paljon henkilötyöpäiviä. Syrjintäperusteita koskevaa tietoa ei saada seurattua. Tarvittava data ei ole valvontatietojärjestelmästä rakenteisena datana, joten raportoinnin yhteydessä tieto etsitään lukemalla asiakastietoja.	Toteutui osittain
<i>Tarvittavat toiminnallisuudet, ominaisuudet ja ulkoiset vaatimukset</i>	Visuaalisen tilannekuvan tulee olla nopeasti päivitettävissä. Visuaalisen tilannekuvan tukena on oltava ristiintaulukoitua numeerista tietoa. Raportoinnissa eikä tilannekuvassa käsitellä henkilötietoja.	Toteutui osittain
<i>Odotettu lopputulos</i>	Raportointityökalu Excel-tiedostona, jossa rakenteinen data kuvataan visuaalisena tilannekuvana. Asiakasyhteydenoton kuvauksesta muodostettu rakenteista dataa. Loppukäyttäjälle ohje raportointityökalun käyttöön.	Toteutui Ei toteutunut Toteutui
<i>Hinta ja aikataulukutus</i>	Aikataulukutus: käytettävissä vuoden 2023 vuosiraportointiin eli tammi-maaliskuun 2024 aikana.	Toteutui

Kehitystyössä hyödynnettävä tutkimusaineisto koostui viranomaisen rekisterimerkinnöistä ja opinnäytetyössä tarkasteltiin tutkimusaineiston datan laatua. Datan laatu koostuu analysoitavuudesta, ajantasaisuudesta, asiaankuuluvuudesta, saatavuudesta, täydellisyydestä, tarkkuudesta ja paikkansapitävyydestä. Kehitystyön aikana datassa laadussa tunnistettiin puutteita. Puutteita havaittiin datan tarkkuudessa ja täydellisyydessä. Osasta yhteydenottoja puuttui tieto toimialasta, työnantajan koosta tai yhteydenoton kuvaus

oli tyhjä. Merkittävin datan laatupuute oli se, ettei asiakasyhteydenoton kuvausteksti ole soveltuvaa käyttötarkoitukseensa syrjäntäperusteiden esiintuomisessa. Näin ollen syrjäntäperusteiden esiintymisfrekvenssit jäävät herkästi hiljaiseksi tiedoksi manuaalisen analyysityön hitauden takia.

Asiakasyhteydenotto on voinut kohdistua useisiin eri asioihin, kuten esimerkiksi häirintään ja epäasialliseen kohteluun, psykososiaaliseen kuormitukseen taikka seksuaaliseen häirintään. Opinnäytetyön tavoitteena oli luoda raportointityökalu työsyryntään liittyville yhteydenotolle. Kehitystyön aikana havaittiin, että löydetty ratkaisu mahdollistaa useamman asialuokan yhteydenottojen tilannekuvan muodostuksen kokonaisuutena. Toimeksiantajalle palautettu raportointityökalu mahdollistaa neljän eri asialuokan raportoinnin. Näkymien erottamiseksi tilannekuvaan lisättiin toiminnallisuutena asialuokkakohtaiset osittajat, jotta toimeksiantaja pystyy tarkastelemaan joko kaikkia neljää asialuokkaa samanaikaisesti tai jokaista erikseen.

Toimeksiantajan tilauksesta työssä tutkittiin erikseen myös sitä, kuka tai kenen suunnasta yhteydenottaja kokee syrjäntää. Koko tutkimusaineistosta löytyi kaiken kaikkiaan yhdeksän yhteydenottoa, joista ilmeni, että syrjäntää koetaan työkaluun suunnasta. Pääasiallisesti syrjäntä koettiin siis työnantajan tai työnantajan työnjohto-oikeutta käyttävän esihenkilön suunnasta.

5.5 Kehitystyö toteutuksen arviointi

Yleisesti tutkimuksen luotettavuus kytkeytyy siihen, että on tutkittu sitä mitä oli tarkoituskin tutkia. (Metsämuuronen, 2006, s. 55) Toimintatutkimuksen luotettavuutta ei pystytä mittaamaan objektiivisin kriteerein. Sen sijaan toimintatutkimuksen luotettavuutta arvioidaan sen mukaan, onko tulokset totta, onko toimintatutkimuksen tulokset ymmärrettäviä, onko esittäjällä perusteet tulosten esittämiselle sekä onko esittäjä ollut vilpitön.

Toimintatutkimuksen luotettavuutta voidaan tutkia myös eri näkökulmista. Esimerkiksi, että kuinka paljon muiden osallistujien kanssa on oltu vuorovaikutuksessa kriittisen ajattelun edistämiseksi tai siitä, millainen muutos on aikaansaatu volyymiltään. Luotettavuutta voidaan arvioida myös tavoitteiden saavuttamisen, tutkimusprosessin noudattamisen taikka osallistujien ymmärryksen syventymisen kautta. (Jantunen ym., 2013, ss. 33–34)

Kokonaisuudessaan kehitystyö toteutettiin suunnitelman mukaisesti. Sisällönanalyysi koostui useista lukukierroksista ja suunnitemallisista merkinnöistä. Toimeksiantajan tarve raportointityökalulle ja sen käyttöä koskevalle ohjeelle oli viimeistään maaliskuun 2024 aikana. Kehitystyö toteutettiin toimeksiantajan määrittämässä aikataulussa. Työn tavoitteet ja toteutus olivat hyvin käytännönläheisiä. Keskustelu raportointiin soveltuvan työkalun

luomisesta ja sisällöstä käynnistyi, kun toimeksiantajan esitti tarpeen nopeammalle analysointityölle. Kehitystyön aikana valmistui raportointityökalu, jolla toimeksiantaja pystyy luomaan työsyrijännän asiakasyhteydenotoista ajantasaisen tilannekuvan nopeasti ja vähin toimenpitein. Toimeksiantajan asiantuntijat osallistuivat raportointityökalun määrittelyihin.

Tämän opinnäytetyön tarkoituksena ei ole ollut tulosten laajempi yleistettävyys, vaan tarkoituksena oli saada syvälinen ja yksityiskohtainen käsitys tutkimusaineistosta eli asiakasyhteydenottojen sisällöstä. Työssä on toteutettu teorialähtöinen temaattinen analyysi. Tutkimusvaiheessa analyysi toteutettiin täsmällisellä tasolla eli luokittelut ja teemoitelu tehtiin ilman tulkintaa ja käytettiin aineistossa täsmällisesti esiintyviä ilmaisuja.

Toimeksiantajan hallussa olevan datan laadun arviointi koettiin tärkeäksi osaksi kehitystyötä, koska kirjallisuuden mukaan laatupuutteiden on todettu hidastavan raportoinnin toteuttamista, puutteet luovat epäselvyyksiä ja pahimmassa tapauksessa päätöksenteko perustuu vääristyneeseen ja epäluotettavaan tietoon (Väre, 2019, ss. 201–202). Laadukas data turvaa sen, että päätökset pohjautuvat relevanttiin, ajantasaiseen sekä luotettavaan tietoon ja kustannustehokkain tapa turvata tämä on huolehtia, että organisaation keräämä ja tallentama data on laadukasta.

Raportointityökalussa työsyrijännän esiintymismuotojen esiintuonnista luovuttiin kehittämistyön tulosten osoittaessa, ettei luotettavia ja yksilöllisesti yhtä syrjintäperustetta kuvaavia hakusanoja pystytty luomaan. Tilannekuvan rooli on tuottaa luotettavaa tietoa päätöksenteon tueksi ja jaettavaksi sidosryhmille, joten hyödynnettävän datan laadun kohdalla on oltava kriittinen. Tutkimushavainnon jälkeen päätettiin, että tilannekuvassa hyödynnetään vain valmiiksi rakenteellisena olevaa dataa. Päätös johti myös siihen, ettei työsyrijintää koskevista asiakasaloitteista saada luotua täysimääräisesti tavoitteena ollut kokonaiskuvaa. Tästä huolimatta kokonaisuutena toimeksiantajan raportointitehokkuus on nyt aiempaa paremmalla tasolla.

6 Pohdintaa ja kehitysehdotuksia

Nyt toteutetun kehitystyön tulokset osoittavat, ettei toimeksiantaja pysty nykyisen datan perusteella luomaan kokonaiskuvaa työsyrijännän eri esiintymismuodoista. Toimeksiantajalla on kuitenkin selkeä tarve työsyrijännän syrjintäperusteiden analysoinnille ja raportoinnille. Pelkkä selkeä tarve ei kuitenkaan johda itsestään soveltuvan datan muodostamiseen, vaan asia vaatii aktiivista tiedon johtamista. Toimeksiantaja on pyrkinyt syrjintäperusteiden yhdenmukaiseen ja säännölliseen keräämiseen, mutta asian loppuunsaattaminen on jäänyt

keskeneräiseksi. Aiemman muutoksen lopputulos on, että työsyRJinnän eri esiintymismuodot ovat edelleen organisaatiossa pääosin hiljaisena tietona.

Raportointityön tehostuminen vaatii soveltuvien tietojärjestelmien, järjestelmäympäristön ja osaamisen lisäksi erityisesti laadukasta dataa sekä tiedon muodostumisen prosessin johtamista. Tietojärjestelmien, järjestelmäympäristön sekä osaamisen osalta toimeksiantajan edellytykset ovat riittävät. Olennaisin tässä tutkimuksessa tunnistettu este raportoinnin tehokkuuden kasvattamiselle on datan laatu puutteet. Toimeksiantajan on jatkossa kiinnitettävä huomiota erityisesti siihen, että se turvaa ilmiön tutkimiseen, analysointiin tai kuvaamiseen soveltuvan datan keräämisen ja luonnin. Tämä vaatii toimeksiantajalta myös vahvempaa tiedolla johtamisen kulttuuria ja tiedon johtamista.

Pääasiallisesti raportoinnin tehostamisen esteeksi todettiin se, ettei syrjäntäperusteita koskeva data ole analysoitavissa. Tälle tunnistettiin kaksi syytä. Ensimmäinen syy liittyy käytettävään tietojärjestelmään. Vaikka työsuojeluviranomaisella on käytössä asiakasyhteydenoton käsittelyyn soveltuva tietojärjestelmä, niin vain osa yhteydenoton tiedoista tallennetaan rakenteellisena datana. Syrjäntäperuste olisi mahdollista tallentaa rakenteellisena datana erillisen tiedonkeruulomakkeen kautta. Tiedonkeruulomake kuitenkin jää usein tallentamatta, koska sen lisääminen jää käyttäjän muistin varaan. Käytännössä käyttöliittymän visuaalinen syöte tiedonkeruulomakkeen lisäämiselle on heikko eli kehote tiedonkeruulomakkeen lisäämismahdollisuudesta ei riittävästi erotu käyttöliittymästä. Lisäksi tiedonkeruulomakkeen tallennusta ei ole määritetty pakolliseksi. Toinen syy liittyy kirjaamiskäytäntöihin. Tämän opinnäytetyön tulokset osoittavat sen, ettei yhteydenoton kuvauksen tekstidatasta pystytä luotettavasti tunnistamaan eri syrjäntäperusteita, koska syrjäntäperusteen kirjaamisessa ei käytetä yhdenmukaista tapausluokittelua. Toteutettu asiakasyhteydenottojen sisällönanalyysi osoittaa siten, ettei nykyiset kirjaamiskäytännöt tuota luotettavaa tietoa toimeksiantajan päätöksenteon tueksi, koska tarvittavaa tietoa ei kerätä määrämukaisesti.

Sisäasiainministeriön vuonna 2009 julkaisemassa tutkimuksessa oli osin käsitelty samankaltaista aineistoa kuin nyt ja tuossa tutkimuksessa annettiin kehitysehdotuksia työsyRJintään liittyvien yhteydenottojen kirjausten sisältöön. Näiden kehitysehdotusten mukaan työsuojeluviranomaisen tulisi kirjata muun muassa syrjäntäperuste yhdenmukaisen tapausluokituksen mukaisesti. Tämän nähtiin nopeuttavan tapauksen luokittelua, kun arkistoitua tapauskuvausta ei tarvitse manuaalisesti etsiä ja lukea erikseen joka kerta. (Aaltonen ym., 2009, s. 25) Sisäasiainministeriön tutkimuksessa esitetty kehitysehdotus yhdenmukaisen tapausluokituksen käytöstä on edelleen täysin relevantti.

Opinnäytetyön kehittämisosuuden tulokset osoittavat, että syrjäntäperusteiden kirjaamisessa on edelleen sellaisia puutteita, joiden takia analysointi on edelleen tehtävä manuaalisesti eikä raportoinnin täysautomatisointi tältä osin ole vielä mahdollista. Toimeksiantaja pystyisi parantamaan datan laatua ainakin kahden eri kokonaisuuden kautta:

1) kirjausohjeistusta tarkentamalla ja 2) toteuttamalla tietojärjestelmämuutoksia.

Kirjausohjeistuksen tarkentaminen toistaa sisäasiainministeriön kehitysehdotusta eli kirjauksissa tulisi käyttää yhdenmukaista tapausluokittelua. Kirjausohjeistuksen tarkentaminen nousee tarpeelliseksi myös siksi, että yhteydenotoista ei pystytty tunnistamaan yhtään syrjäntätapausta 30 %:ssa tapauksia.

Toimeksiantajan on hyvä ottaa käyttöön jokaisessa kirjauksessa sellainen vakiomuotoinen tapausluokittelu, joka mahdollistaa tehokkaan analysointityön. Käytännössä kirjaamiskäytännön muutos voisi olla sellainen, että kuvastekstin loppuun lisätään syrjäntäperusteet yhdenmukaisen luokittelun mukaisesti esimerkiksi kirjaamalla selkeästi yksilöivä syrjäntäperuste (ikä, alkuperä, kansalaisuus, kieli uskonto, vakaumus, mielipide, poliittinen toiminta, ammattiyhdistystoiminta, perhesuhteet, terveydentila, vammaisuus, seksuaalinen suuntautuminen, muu henkilöön liittyä syy). Kirjaamiskäytäntöjen muutos on nopeammin toteutettavissa kuin tietojärjestelmämuutos. Kirjaamiskäytäntöjen muutos vaatii toimeksiantajalta jatkuvaa ja tehokasta seuranta datan korkean laadun varmistamiseksi.

Toinen vaihtoehtoinen tapa huolehtia siitä, että syrjäntäperuste tulee tallennettua rakenteellisena datana, on toteuttaa muutos tietojärjestelmään. Toimeksiantajan tietojärjestelmässä on jo nyt mahdollista erillisen tiedonkeruulomakkeen liittäminen. Kun tietoa kerätään ja tallennetaan tiedonkeruulomakkeen avulla, tulisi tiedonkeruulomakkeen lisääminen olla helppoa ja jopa automaattista, ja tallentamisen tulisi olla pakollista. Tietojärjestelmämuutos mahdollistaisi syrjäntäperusteiden tallentamisen rakenteellisena data, mikä maksimoisi analysoinnin tehokkuutta, kun analysointityö voitaisiin jopa automatisoida. Toimeksiantaja voisi ottaa käyttöönsä myös muita tietojärjestelmäratkaisuja kuten esimerkiksi laajentaa työsuojeluviranomaisen digitaalista asiointipalvelua, jonka kautta työntekijät voivat jo nyt tehdä valvontapyyntöjä muista asioista (Työsuojelu.fi, 2021, 2022).

Mikäli toimeksiantaja haluaa siirtyä tekoälyn hyödyntämiseen esimerkiksi ohjatun koneoppimisen keinoin, on viranomaisen rekistereihin tallennettavan datan laatuun kiinnitettävä huomiota. Kaikissa tekoälyyn liittyvissä ohjelmistoissa korostuu datan hyödyntäminen, koska dataa hyödynnetään koneoppimismallien sisältämien algoritmien opettamisessa. (Merilehto, 2018, ss. 27–44)

Tekoälyratkaisujen hyödyntämisen mahdollisuus nousi esiin opinnäytetyön ennakkokeskusteluissa. Alkuperäinen tutkimussuunnitelma sisälsi tämän näkökulman, mutta siitä luovuttiin aluehallintoviraston tekoälylinjausten julkaisun jälkeen. Tekoälyratkaisuihin voidaan lukea niin luonnollisen kielen käsittelyyn pohjautuvien generatiivisten tekoälymallien lisäksi algoritmiperusteiset ratkaisut. Luonnollisen kielen käsittelyyn liittyvien tekoälysovellusten hyödyntäminen asiakasyhteydenottojen kuvausten analysoinnissa on tällä hetkellä mahdotonta, koska salassa pidettäviä tietoja ei nykyisen tekoälylinjauksen mukaan saa käsitellä tekoälysovelluksilla. Generatiivisen tekoälyn käyttöönotto vaatii toimeksiantajalla muun muassa suljettua järjestelmäympäristöä, josta analysoitava aineisto ei siirry järjestelmän ulkopuolelle analysointivaiheessa eikä edes kielimallia päivitettäessä. Tekoälyä, algoritmiperusteisia ratkaisuja tai koneoppimISRatkaisuja ei pystytä hyödyntämään salassa pidettävän aineiston analysointiin ennen kuin toimeksiantajalta on siihen soveltuva oma suljettu ympäristö ja järjestelmä.

Vaikka tämän tutkimuksen toteutuksessa luovuttiin tekoälyn hyödyntämisestä, niin tekoäly on todennäköistä tulevaisuutta myös työsuojeluhallinnossa. Tekoälyn taustalta löytyy tilastotiedettä, datatiedettä ja ohjelmointia. Tekoäly poikkeaa sääntöpohjaisesta ohjelmoinnista siten, ettei tekoälylle anneta valmiiksi sääntöjä tai ohjeita algoritmien muodossa vaan sen sijaan tekoälyohjelma tai -sovellus tunnistaa aineistossa toistuvat säännönmukaisuudet. (Kananen & Puolitaival, 2019, ss. 27–32)

Tekoälyn voidaan katsoa jakautuvan vakiintuneen käytännön mukaan koneoppimiseen sekä neuroverkkoihin ja syväoppimiseen (Merilehto, 2018, ss. 17–21). Näen mahdolliseksi sen, että tulevaisuudessa myös salassa pidettäviä aineistoja tutkittaisiin koneoppimisen menetelmin sen jälkeen, kun analysointiin soveltuva ympäristö on saatu käyttöön. Työsyryjäntäilmiöön soveltunee parhaiten ohjattu koneoppiminen, jossa mallia opetetaan aineistolla, joka sisältää syötteen lisäksi lopputuloksen (Merilehto, 2018, ss. 17–19). Tämän opinnäytetyön tulokset osoittavat sen, ettei asiakasyhteydenotot sovellu mallin opetusaineistoksi, koska yhteydenoton kirjaukset eivät sisällä tietoa vasteesta eli lopputuloksesta. Tämä käy ilmi opinnäytetyön tutkimusaineistosta, jossa suureen osaan ilmoituksista oli merkitty, että yhteydenotto ei johda edes valvonnallisiin toimenpiteisiin. Ohjatun koneoppimismallin opetusdatan tulee siis sisältää tieto yhteydenoton kirjauksesta sekä suoritettujen valvonnan lopputuloksesta.

Tässä opinnäytetyössä analysointityö toteutettiin manuaalisesti Excelillä, koska toimeksiantajalla ei ole käytössä muita ohjelmia tai teknisiä ratkaisuja kuvaustekstien analysointityöhön. Hakusanoja luotiin tyypistämällä, koska toimeksiantajalla ei ole käytössä tekniikoita, joilla voitaisiin toteuttaa vapaan tekstin tyypistämistä eli stemmausta. Stemmaus

on typistämistekniikka, jolla eri sijamuodoissa tai sanaluokissa olevista sanoista tunnistetaan yhteinen juurimuoto (Harjunpää, 2004, ss. 5–6). Tulevaisuudessa toimeksiantajan on hyödyllistä tutustua myös jo olemassa oleviin algoritmeihin, joiden avulla vapaata tekstiä pystyy paremmin analysoimaan. Esimerkiksi Martin Porterin luoma tekstin stemmaukseen soveltuva Snowball-algoritmi, josta löytyy myös suomen kielelle soveltuva versio (Porter, n.d.).

7 Johtopäätökset

Tämän opinnäytetyön tutkimuskysymyksiä käsiteltiin toimeksiantajan hallussa olevan datan kautta. Tutkimuskysymykset olivat:

1. Miten toimeksiantajan raportointitehokkuutta voidaan parantaa?
2. Kuinka työsyöjinnän kokonaistilanne saatiin selvitettyä?

Työn aikana toimeksiantajan raportointitehokkuutta parannettiin luomalla raportointityökalu. Työn aikana havaittiin toimeksiantajan datan laadussa puutteita, joiden takia raportointitehokkuutta pystyttiin tehostamaan vain rakenteellisen datan osalta. Datan laatu puutteille tunnistettiin syitä, joiden perusteella toimeksiantajalle esitettiin kehitysehdotuksia tilanteen korjaamiseksi. Tämän lisäksi työn aikana muodostui kokonaiskuva työsyöjintätilanteista, joista oli oltu yhteydessä työsuojeluviranomaiseen vuonna 2023.

Työhön valittu näkökulma tehostamiseen oli raportointiin kuluvan ajan lyhentäminen.

Konkreettiseksi tehostamiseksi valikoitui raportointityökalun kehittäminen.

Toimeksiantajalle luotu raportointityökalu auttaa toimeksiantajaa tuottamaan ajantasaista tilastotietoa nykyistä nopeammin ja vähäisemmällä työllä. Toimeksiantajalla on nyt mahdollisuus luoda tilannekuva nopeasti ja vähin toimenpitein työsyöjinnän lisäksi kolmen muun asialuokan asiakasyhteydenotoista. Toimeksiantaja pystyy luomaan tilannekuvan raportointityökalulla niin usein kuin on tarpeen. Näiltä osin toimeksiantajan raportointityö on tehostunut.

Raportointityökalu ja visuaalinen tilannekuva työsyöjinnän asiakasyhteydenotoista parantavat asiakasyhteydenottojen hyödyntämismahdollisuuksia tiedolla johtamisen näkökulmasta.

Toimeksiantajalle luotu raportointityökalu ja tilannekuva parantavat näin toimeksiantajan kykyä hyödyntää asiakasyhteydenottoja oman toiminnan suunnittelussa, viestinnässä ja tilastoinnissa. Samoin tietopyyntöihin vastaaminen nopeutuu.

Toimeksiantaja on ottanut valmistuneen raportointityökalun käyttöön. Työkalu mahdollistaa myös trendimuutosten havaitsemisen nopeammin ja tästä toimeksiantajalle on kertynyt kokemusta raportointityökalun lyhyen käyttöhistorian aikana. Työsuojeluviranomaisella ja yhdenvertaisuusvaltuutetulla on jaettu toimivalta syrjintään liittyvissä asioissa 1.6.2023 alkaen. Toimeksiantaja oli ennakoinut, että toimivallan jakautumisen jälkeen heille saapuisi vähemmän työsyryntään liittyviä asiakasyhteydenottoja. Raportointityökalua hyödyntämällä toimeksiantaja totesi, ettei toimivallan jakautuminen ole johtanut asiakasyhteydenottojen vähenemiseen.

Työ paljasti toimeksiantajan datan laadussa puutteita. Olennaisin todettu laatupuute on, ettei asiakasyhteydenoton kuvaustekstistä pystytä tunnistamaan työsyryntään esiintymismuotoja eli syrjintäperusteita. Kerätty data ei siten sovellu analysointityöhön. Datan laatu on kriittisessä asemassa myös tekoälyratkaisujen kehitystyössä. On syytä todeta, ettei tutkimusaineisto sisällä kaikkia toimeksiantajan järjestelmään tallennettuja tietoja yhteydenotoista. Toimeksiantajan datassa voi olla muitakin puutteita, joita ei ole pystytty havainnoimaan. Toimeksiantajan on hyvä toteuttaa kokonaisvaltainen datan laatutekijöiden arviointi sekä juurisyyanalyysi havaitsemilleen laatupuutteille.

Yleiskielen tiedolla on viisi tasoa: data, informaatio, tieto, näkemys, viisaus. Päätöksenteon tueksi tarvittava viisaus ei synny tyhjiöstä. Päätöksenteon tueksi tarvitaan viisautta, joka on muodostettu relevantin ja oikeassa kontekstissa olevasta ymmärryksestä, tietämyksestä, informaatiosta ja viime kädessä datasta. Tiedon jalostumisen prosessi eli viisauden muodostumisen vaatii yhtä lailla aktiivista johtamista ja vuorovaikutusta. Tietojohtamisen ideaalitalanteessa päätöksenteon tarpeet ohjaavat tiedon muodostumista. Päätöksenteon tietotarpeiden tulee siten ohjata toiminnan aikana kerättävää tietoa jopa yksityiskohtaisen datan tasolla. Tämän voisi ilmaista myös siten, että jos data ei tuota päätöksenteossa tarvittavaa tietämystä niin syynä ei ole datan puute. Syynä voi olla tiedolla johtamisen kulttuurin tai strategian puute taikka tehottomat tiedolla johtamisen prosessit, jotka eivät riittävästi johda tiedon keräämistä.

Toimeksiantajalle oli esitetty kehitysehdotuksia asiakasyhteydenottojen kirjausten sisällöstä jo vuonna 2009. Tämän kehitystyön aikana ilmeni, että aiempien kehitysehdotusten perusteella tehdyt toimenpiteet ovat jääneet keskeneräiseksi. Toimeksiantajan on hyvä jatkaa kehitystoimenpiteitä. Jatkossa toimeksiantajan on kiinnitettävä huomiota erityisesti tiedon johtamiseen eli siihen, että se turvaa ilmiön tutkimiseen, analysointiin tai kuvaamiseen soveltuvan datan keräämisen ja luonnin. Tämä vaatii toimeksiantajalta myös vahvempaa tiedolla johtamisen kulttuuria ja tiedon johtamista.

Tämän opinnäytetyön pääasiallinen kehitysehdotus kohdistuu tiedon johtamiseen ja erityisesti kirjauskäytäntöjen yhdenmukaistamiseen. Syrjintäperusteiden esiintymisfrekvenssit saadaan esille yhdenmukaisen tapausluokittelun käyttöönotolla.

Loppuyhteenvetona todettakoon, että toimeksiantajan raportoinnin toteuttamista onnistuttiin tehostamaan. Toimeksiantaja katsoo hyötывänsä toteutetusta kehitystyöstä ja sisällönanalyysistä, vaikka tilannekuvan sisällössä tehtiin kompromisseja. Toimeksiantaja hyötyy ajantasaisesta raportoinnista siten, että päätöksenteko perustuu ajantasaiseen ja luotettavaan tilannekuvaan. Toimeksiantajalle on nyt luotu raportointia varten työkalu, jolla saa luotua tilannekuvan nopeasti ja tilannekuva on skaalattavissa myös muihin asiakasyhteydenottojen asialuokkiin kuin työsyrijintään. Kehitysehdotusten käytäntöön viennin jälkeen toimeksiantajalla on edellytyksiä tehostaa lisää raportointiaan.

Lähteet

- Aaltonen, M., Joronen, M., & Villa, S. (2009). Syrjintä työelämässä – pilottitutkimus työsuojelupiirien aineistosta. <http://urn.fi/URN:ISBN:978-952-491-548-9>
- Aihkisalo, Halén, M., Holmström, H., Jurmu, P., Matinmikko, T., Seppälä, T., Tihinen, M., & Tirronen, J. (2018). Ohjelmistorobotiikka ja tekoäly – soveltamisen askelmerkkejä. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja, 65/2018.
- Ailinpieti, H., & Ruottinen, J. (2002a). MVP Canvas. Oulun yliopisto. <https://aoe.fi/api/v1/download/file/mvpcanvasnettiin-1655285553102.pdf>
- Ailinpieti, H., & Ruottinen, J. (2002b). MVP Canvas Ohjeistus. Oulun yliopisto. <https://aoe.fi/api/v1/download/file/mvppvaiheistus03docx1-1655285553132.pdf>
- Aluehallintovirasto. (2023). Aluehallintovirastojen tekoälylinjaukset kannustavat kokeiluihin turvallisesti ja vastuullisesti, Tiedote 11.12.2023. <https://avi.fi/tiedote/-/tiedote/70070269>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Cahyadi, A., & Prananto, A. (2015). Reflecting design thinking: A case study of the process of designing dashboards. *Journal of Systems and Information Technology*, 17(3), 286–306. <https://doi.org/10.1108/JSIT-03-2015-0018>
- Cong, X., & Pandya, K. V. (2003). Issues of Knowledge Management in the Public Sector. 1(2). https://static.aminer.org/pdf/PDF/000/243/343/knowledge_management_issues_challenges_and_opportunities_for_governments_in_the.pdf
- Elomaa-Krapu, M. (n.d.). TEMAATTINEN ANALYYSI. <https://events.tuni.fi/uploads/2019/09/ffe0611d-mita-on-temaattinen-analyysi.pdf>
- Eskelinen, T., Harjunen, H., Hirvonen, H., & Jokinen, E. (Toim.). (2017). Tehostamistalous. <https://jyx.jyu.fi/handle/123456789/53297>

Eskola, J., Lähti, J., & Vastamäki, J. (2018). Teemahaastattelu: Lyhyt selviytymisopas.

Teoksessa R. Valli (Toim.), Ikkunoita tutkimusmetodeihin 1. Metodien valinta ja aineistonkeruu: Virikkeitä aloittavalle tutkijalle (5. uudistettu painos). PS-Kustannus.

Finto Suomalainen asiasanasto- ja ontologiapalvelu. (n.d.a). Rakenteeton data.

<http://urn.fi/URN:NBN:fi:au:tt:t53>

Finto Suomalainen asiasanasto- ja ontologiapalvelu. (n.d.c.). Rakenteinen data.

<https://finto.fi/tt/fi/page/t98>

Finto Suomalainen asiasanasto- ja ontologiapalvelu. (n.d.b). Tieto.

<http://urn.fi/URN:NBN:fi:au:tt:t117>

Gil, Y., Khider, D., Osorio, M., Ratnakar, V., Vargas, H., & Garijo, D. (2022). Towards Capturing Scientific Reasoning to Automate Data Analysis. Proceedings of the Annual Conference of the Cognitive Science Society, 44.

Harjunpää, M. (2004). Tekstidokumenttien klusterointi. Teoksessa E. Mäkinen (Toim.), Pieniä tietojenkäsittely- tieteellisiä tutkimuksia (syksy 2003) (ss. 1–13).

https://webpages.tuni.fi/utacs_history/cs/reports/bsarja/B-2004-2.pdf

Hyvärinen, M., Suoninen, E., & Vuori, J. (n.d.). Haastattelut. Teoksessa J. Vuori (Toim.), Tietoarkisto. <https://www.fsd.tuni.fi/fi/palvelut/metodit/laadullisen-tutkimuksen-aineistot/haastattelut/>

Jantunen, S., Naaranoja, M., Piippo, J., Mäkelä, T., Valtanen, E., & Sankel, M. (2013).

Toimintatutkimuksen laatukriteerit ja niiden soveltaminen osallistuvan innovaatiotoiminnan johtamisen tutkimisessa. Työelämän tutkimuspäivät (2012), 29–56.

Jauhola, L., Siltala, J., & Nieminen, K. (2022). ”Että puututtaisiin konkreettisesti”

Seurantaselvitys vihapuheesta ja häirinnästä ja niiden vaikutuksista eri vähemmistöryhmiin. <https://urn.fi/URN:ISBN:978-952-400-482-4>

Juhila, K. (n.d.a). Koodaaminen. Teoksessa J. Vuori (Toim.), Laadullisen tutkimuksen verkkokäsikirja.

<https://www.fsd.tuni.fi/fi/palvelut/metodit/laadullisen-tutkimuksen-aineistot/haastattelut/>

- Juhila, K. (n.d.b). Laadullinen tutkimus ja teoria. Teoksessa J. Vuori (Toim.), Laadullisen tutkimuksen verkkokäsikirja.
<https://www.fsd.tuni.fi/fi/palvelut/metodit/metodit/kvali/mita-on-laadullinen-tutkimus/laadullinen-tutkimus-ja-teoria/>
- Jyrkämä, J. (n.d.). Toimintatutkimus. Teoksessa J. Vuori (Toim.), Laadullisen tutkimuksen verkkokäsikirja.
<https://www.fsd.tuni.fi/fi/palvelut/metodit/metodit/kvali/tutkimusasetelma/toimintatutkimus/>
- Kanninen, O., Virkola, T., Lilja, E., & Rask, S. (2022). Tavoitteena syrjimätön työelämä – Työsyrjinnän nykytila ja keinoja tasa-arvon ja yhdenvertaisuuden edistämiseksi.
<http://urn.fi/URN:ISBN:978-952-383-362-3>
- Kuusisto, R. (2005). Tilannekuvasta täsmäjohtamiseen. Johtamisen tietovirrat kriisin hallinnan verkostossa. Liikenne ja viestintäministeriö.
https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/78700/Julkaisu_81_2005.pdf?sequence=1
- 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, T. (2013). Tietojohdaminen. Tampereen teknillinen yliopisto Tiedonhallinnan ja logistiikan laitos. <https://urn.fi/URN:ISBN:978-952-15-3058-6>
- Laki naisten ja miesten välisestä tasa-arvosta (609/1986).
<https://www.finlex.fi/fi/laki/ajantasa/1986/19860609>
- Laki työsuojelun valvonnasta ja työpaikan työsuojeluyhteistoiminnasta (44/2006).
<https://www.finlex.fi/fi/laki/ajantasa/2006/20060044>
- Leskelä, R.-L., Haavisto, I., Jääskeläinen, A., Helander, N., Sillanpää, V., Laasonen, V., Ranta, T., & Torkki, P. (2019). Tietojohdaminen ja sen kehittäminen: Tietojohdamisen arviointimalli ja suosituksia maakuntavalmistelu pohjalta. 2019:42.
<http://urn.fi/URN:ISBN:978-952-287-754-3>
- Listenmaa, J. (2023). Laita tieto töihin. Tiedolla johtamisen käsikirja. Alma Talent Oy.

- Metsämuuronen, J. (2006). Tutkimuksen tekemisen perusteet ihmistieteissä. International Methelp Ky.
- Mykkänen, M., Ahonen, O., Kinnunen, U.-M., & Saranto, K. (2023). Rakenteisesti kirjatun hoitotyön tiedon vuodeosastotason analyysi – esimerkki FinCC-luokituskokonaisuuden käytöstä. *Finnish Journal of EHealth and EWelfare*, 15(2), 142–156. <https://doi.org/10.23996/fjhw.109940>
- Mykkänen, M., Miettinen Merja, Siponen, T., & Saranto, K. (2021). Sairaalan reaaliaikainen tilannekuva päivittäisessä johtamisessa. *Finnish Journal of EHealth and EWelfare*, 13(4), 425–441.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria—Lorelli S. Nowell, Jill M. Norris, Deborah E. White, Nancy J. Moules, 2017. *International Journal of Qualitative Methods*, 16(1), 1–13.
- Pietiläinen, M., Viitasalo, N., Lipiäinen, L., Ojala, S., Leinonen, M., Otonkorpi-Lehtoranta, K., Jokinen, E., Korvajärvi, P., & Nätti, J. (2018). TYÖSSÄ KOETTU SYRJINTÄ JA MYÖHEMPI TYÖURA Loppuraportti. Tampereen yliopisto. <https://trepo.tuni.fi/bitstream/handle/10024/102975/978-952-03-0662-5.pdf?sequence=1>
- Rikoslaki (39/1889). <https://www.finlex.fi/fi/laki/ajantasa/1889/18890039001>
- Saaranen-Kauppinen, A., Puusniekka, A., Kuula, A., Rissanen, R., & Karvinen, I. (2009). Menetelmäopetuksen tietovaranto KvaliMOTV. Kvalitatiivisten menetelmien verkkopöytäkirja. Tampereen yliopisto Yhteiskuntatieteellisen tietoarkisto. <https://www.fsd.tuni.fi/fi/tietoarkisto/julkaisut/kvalimotv.pdf>
- Salminen, V., Halme, K., Järvelin, A.-M., Kettinen, J., Uusikylä, P., Lintinen, U., Stenvall, J., Vakkuri, J., & Johanson, J.-E. (2021). Valtionhallinnon tulosohtausmallin arviointi. <http://urn.fi/URN:ISBN:978-952-383-425-5>
- Sanastokeskus. (2023). Tilannekuva. <https://termipankki.fi/tepa/fi/haku/tilannekuva>
- SFS-ISO/IEC 25012:2020. (2020). Software engineering. Software product Quality Requirements and Evaluation (SQuaRE). Data quality model. SFS Online.

- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286.
<https://doi.org/10.1016/j.jbusres.2016.08.001>
- Sivula, A., Aho, M., & Laukkanen, M. (2023). *Datasta liiketoimintaan: 10 tehokasta työkalua*. Alma Talent Oy.
<https://ezproxy.hamk.fi/login?url=https://bisneskirjasto.almatalent.fi/teos/23bi448454>
- Sosiaali- ja terveystieteiden ministeriö. (2016). *TYÖSUOJELUVALVONNAN UUSI TAPA TARKASTAA – VALTIMO-HANKE* Valvonnan parempaa vaikuttavuutta paremmalla tiedonhallinnalla. <http://urn.fi/URN:ISBN:978-952-00-3767-3>
- Takala, P. (2018). *Kohti jatkuvaa tilannekuvaa*.
<https://media.sitra.fi/app/uploads/2018/11/kohti-jatkuvaa-tilannekuvaa.pdf>
- Tilastokeskus. (n.d.). *Validiteetti*. <https://www.stat.fi/meta/kas/validiteetti.html>
- Työsopimuslaki (55/2001). <https://www.finlex.fi/fi/laki/ajantasa/2001/20010055>
- Työsuojelu.fi. (2021). *Työsuojeluhallinnon digitaalisen asiointipalvelun pilotti on käynnistynyt*.
Työsuojelu. <https://tyosuojelu.fi/-/tyosuojeluhallinnon-digitaalisen-asiointipalvelun-pilotti-on-kaynnistynyt>
- Työsuojelu.fi. (2022). *Työsuojelun asiointipalvelua laajennetaan*. Työsuojelu.
<https://tyosuojelu.fi/-/tyosuojelun-asiointipalvelua-laajennetaan>
- Työsuojeluhallinto. (2023). *Yhdenvertaisuuden ja syrjinnän kiellon valvonta työelämässä vuonna 2022*.
https://tyosuojelu.fi/documents/154017715/168016404/TSH_Raportti_Syrjinnanvalvonta2022.pdf/02f01cf2-106f-ce57-25f1-f50af5023b92/TSH_Raportti_Syrjinnanvalvonta2022.pdf?t=1684915270403
- Työturvallisuuslaki (738/2002). <https://www.finlex.fi/fi/laki/ajantasa/2002/20020738>
- Valtionvarainministeriö. (2019). *Taloushallinnon automaation ja raportoinnin kehittämisen koordinaatioryhmän loppuraportti*. https://api.hankeikkuna.fi/asiakirjat/27d03df3-5565-4eec-838c-64458464ba81/36408df7-65d7-4e04-bce9-ed0f9845a730/JULKAISU_20190411132000.PDF
- Väre, T. (2019). *Master data*. Alma Talent Oy.

Yhdenvertaisuuslaki (1325/2014). <https://www.finlex.fi/fi/laki/alkup/2014/20141325>

Liite 1. Aineistonhallintasuunnitelma

Aineistonhallintasuunnitelma

Päivitetty 22.12.2023

Tutkimusaineisto, sen tallennus ja säilytys

Opinnäytetyötä varten ei kerätä erikseen tutkimusaineistoa. Tutkimusaineisto koostuu viranomaisen hallussa olevista asiakasyhteydenotoista. Opinnäytetyötä varten asiakasyhteydenotoista pyydetään seuraavat tiedot: ilmoituksen vastaanoton ajankohta, ilmoittajan asema tai rooli organisaatiossa, yhteydenoton lähde eli ilmoituskanava, yhteydenoton asialuokat, toimiala, yhteydenoton kuvaus, yhteydenoton järjestelmään merkityt jatkotoimenpiteet sekä arvioitu toimenpide, vastauksen antaminen ilmoittajalle, vastauksen ajankohta sekä ilmoituksen kohteena olleen työnantajan henkilöstön suuruusluokka. Opinnäytetyötä varten ei pyydetä ilmoituksen jättäneen asiakkaan yksilöintitietoja eikä yksilöintitietoja yrityksestä tai työpaikasta, jota ilmoitus koski.

Ennen aineiston luovutusta tutkimusta varten viranomainen on varmistanut, ettei luovutettavat tiedot sisällä ilmoituksen tehneen henkilön yksilöintitietoja eikä tietoja ilmoituksen kohteena olevasta yrityksestä tai työpaikasta. Ennen aineiston luovutusta viranomainen tarvittaessa anonymisoi tiedot siten, ettei tutkittavaa aineistoa pystytä yhdistämään ilmoittajan henkilöllisyyteen, ilmoituksen kohteena olevaan työpaikkaan eikä valvontatoimenpiteeseen. Opinnäytetyösopimukseen on kirjattu, että tutkimusaineisto on anonymisoitava.

Tutkimus toteutetaan opinnäytetyön toimeksiantajan laitteilla. Aineistoa säilytetään ja käsitellään erillisessä järjestelmäkansiossa, johon vain opinnäytetyön tekijällä on käyttöoikeus. Tutkimusaineistoa ei siirretä pois opinnäytetyön tilaajan laitteista eikä aineistoa säilytetä pilvipalvelussa. Kun tutkimusaineistoa säilytetään ja käsitellään vain työn tilaajan laitteilla ja verkossa, on aineisto myös suojattu työn tilaajan eri tietoturvakäytännöillä, kuten käyttäjätunnuksin, salasanoin sekä vahvalla tunnistautumisella järjestelmään. Samalla tietojen varmuuskopiointi tapahtuu automaattisesti joka päivä. Vikatilanteissa tutkimusaineisto voidaan koota uudestaan.

Tutkimuslupa

Tutkimuslupa on haettu Sosiaali- ja Terveysministeriöstä.

Henkilötietojen käsittely opinnäytetyössä

Vera valvontatietojärjestelmän tietosuojaselosteessa todetaan, että viranomainen käsittelee henkilötietoja asiakasyhteydenottojen rekisteröinnissä, seurannassa ja valvontatoimenpiteen toteuttamisessa ja henkilötietojen käsittely viranomaisessa perustuu lakiin. Henkilötietoina kerätään tietoja mm. työpaikan työsuojeluhenkilöstöstä, tarkastuksiin osallistuneista henkilöistä sekä työsuojeluviranomaiseen yhteydessä olleesta asiakkaasta.

Tapauskohtaisesti kerättäviä tietoja on henkilön nimi, yhteystiedot, syntymäaika, ammattinimike ja organisaatio. Työsuojeluviranomaiseen yhteydessä olleen henkilön tiedot ovat työsuojelun valvontalain (44/2006) 10 §:n perusteella salassa pidettäviä. (Länsi- ja Sisä-Suomen aluehallintovirasto, 2023)

Henkilötiedolla tarkoitetaan kaikkia luonnolliseen henkilöön liitettäviä tietoja. Esimerkkejä suorista tunnistetiedoista: nimi, osoite tai henkilötunnus. Esimerkkejä epäsuorista tunnistetiedoista: ammattinimike tai asema, työpaikka, sukupuoli. (Vastuullinen tiede, 2019)

Tutkimusaineisto ei sisällä suoria henkilötietoja. Tutkimusaineisto sisältää epäsuoria henkilötietoja, kuten asema työpaikalla tai sukupuoli. Tutkimusaineisto sisältää tietoja erityisistä henkilötiedoista kuten rotu, etninen alkuperä, poliittinen mielipide, uskonnollinen tai filosofinen vakaumus, ammattiliiton jäsenyys, terveydentilatiedot, seksuaalinen suuntautuminen tai käyttäytyminen, jotka on tallennettu viranomaisen rekisteriin yleisellä tasolla työsuojelutapauksissa. (Tieto- ja viestintävirasto, n.d., Länsi- ja Sisä-Suomen aluehallintovirasto, 2023). Epäsuorat tai erityiset henkilötiedot on esitetty tutkimusaineistossa siten, ettei niitä voi yhdistää ilmoituksen tekijään, ilmoituksen kohteena olevaan työpaikkaan eikä valvontatoimenpiteeseen. Tutkimusaineisto ei sisällä geneettisiä ja biometrisia tietoja henkilön tunnistamista varten.

Opinnäytetyöaineiston omistajuus

Opinnäytetyön aineiston ja tulokset omistaa työn tilaaja.

Opinnäytetyöaineiston jatkokäyttö työn valmistumisen jälkeen

Tutkimusaineistoa ei luovuteta jatkokäyttöön.

Viranomaisen asiakirjojen säilytysajoista on saatavilla lisätietoa Työsuojelun internetsivuilta. Opinnäytetyön ja tutkimusaineiston välitallennusversiot yms. säilytetään opinnäytetyön hyväksymisestä yhden (1) vuoden ajan, jonka jälkeen ne hävitetään.

Lähteet:

Länsi- ja Sisä-Suomen aluehallintovirasto. (2023). Tietosuojaseloste. Vera-
valvontatietojärjestelmä. <https://www.tyosuojelu.fi/tietoa-meista/tietosuoja>

Tietosuojavaltuutetun toimisto. (n.d.) *Erityisten henkilötietojen käsittely*.
<https://tietosuoja.fi/erityisten-henkilotietoryhmien-kasittely>

Vastuullinen tiede. (2019). Käsitletkö kuitenkin henkilötietoja?
<https://vastuullinentiede.fi/fi/tutkimustyo/kasitletko-kuitenkin-henkilotietoja>