

**TEKOÄLYN MAHDOLLISUUKSIEN HYÖDYNTÄMINEN
HENKILÖSTÖTUTKIMUKSEN AVOVASTAUSTEN ANALYSOINNISSA**



Ammattikorkeakoulututkinnon opinnäytetyö

Tietojenkäsittelyn koulutus, Hämeenlinnan korkeakoulukeskus
kevät, 2024

Janica Tarén

Tietojenkäsittelyn koulutus

Tiivistelmä

Tekijä Janica Tarén

Vuosi 2024

Työn nimi Tekoälyn mahdollisuuksien hyödyntäminen henkilöstötutkimuksen avovastausten analysoinnissa

Ohjaajat Lasse Seppänen

TIIVISTELMÄ

Opinnäytetyön aiheena on tekoälyn hyödyntäminen henkilöstötutkimuksen avointen tekstivastausten käsittelyssä. Opinnäytetyön tarkoituksena oli selvittää, millaisia tarpeita valtion organisaatioilla on henkilöstötutkimuksen avovastausten käsittelyprosessiin liittyen ja miten tekoälyn avulla tarpeisiin on mahdollista vastata. Lisäksi tavoitteena oli selvittää, millaisia asiakkaiden tarpeisiin vastaavia ratkaisuja järjestelmän toimittajilla on saatavilla. Opinnäytetyön toimeksiantaja oli Valtion talous- ja henkilöstöhallinnon palvelukeskus.

Opinnäytetyö on tutkimuksellinen. Opinnäytetyön tietopohja koostuu henkilöstötutkimuksen perusteista ja käyttötarkoituksista sekä tekoälyn perusteista ja käyttökohteista etenkin luonnollisen kielen käsittelyn alueella. Opinnäytetyön tutkimusaineisto kerättiin kyselylomakkeilla ja haastatteluilla. Aineisto oli pääasiassa laadullista ja se analysointiin laadullisen analyysin menetelmin. Verkkokysely sisälsi myös määrällistä aineistoa, joka on analysoitu kaavioiksi.

Opinnäytetyössä havaittiin, että avovastausten analysoinnin tarpeisiin on mahdollisuuksia vastata tekoälyn avulla. Opinnäytetyön perusteella koostettiin tekoälyn käyttöönottoon liittyviä huomioita ja hyötyjä, jotka on hyvä huomioida avovastausten analysoinnin ratkaisun kehittämisessä. Opinnäytetyön toimeksiantaja oli tyytyväinen tutkimuksessa esille tulleisiin tietoihin ja lopputulokseen.

Avainsanat Henkilöstötutkimus, luonnollisen kielen käsittely, tekoäly

Sivut 48 sivua ja liitteitä 3 sivua

Degree Programme in Business Information Technology

Abstract

Author Janica Tarén

Year 2024

Subject Possibilities of AI in analysis of a personnel survey's open responses

Supervisors Lasse Seppänen

ABSTRACT

The subject of the thesis is how to utilize artificial intelligence in a personnel survey's open responses. The purpose of the thesis was to find out what kind of needs state organizations have in analyzing open responses and how artificial intelligence may respond to those needs. In addition, the purpose was to find out what kind of solutions system providers can offer, that will meet the needs of the customers. The thesis was commissioned by The Finnish Government Shared Services Centre for Finance and HR.

The thesis is theoretical. The theoretical base includes basic knowledge of personnel surveys and their purposes. Also, the basics and the use of artificial intelligence is discussed especially around natural language processing. The research material of the thesis was collected by surveys and interviews. The data was mainly qualitative and was analyzed using qualitative analysis methods. The online survey also included quantitative data, which has been analyzed into graphs.

The analysis indicates that artificial intelligence has opportunities to meet the needs of analyzing a personnel survey's open responses. Based on the analysis, considerations and benefits of the implementation of artificial intelligence were gathered. It is recommended that these things are considered when the solution of open responses analysis will be developed. The commissioner of the thesis was satisfied with the results of the research.

Keywords Personnel survey, natural language processing, AI

Pages 48 pages and appendices 3 pages

Sanasto

AI	Tekoäly; ihmisen toimintaa ja päätöksentekoa jäljittelevä ohjelmisto.
Algoritmi	Ohje ongelman ratkaisemiseksi.
CNN	Konvoluutioneuroverkko; syväoppimisen neuroverkkotyyppi.
GDPR	EU:n tietosuoja-asetus henkilötietojen suojaamiseksi.
Klusterointi	Datan jakaminen ennalta määrittelemättömiin ryhmiin
NLP	Luonnollisen kielen käsittely; tekniikka, jolla kone käsittelee ihmisen tuottamaa kieltä.
NLU	Luonnollisen kielen ymmärtäminen; tekniikka, jolla kone ymmärtää ihmisen tuottamaa kieltä.
RNN	Takaisinkytketty neuroverkko; syväoppimisen neuroverkkotyyppi.
Sanapilvi	Raportti aineistossa eniten toistuvista sanoista.

Sisälllys

1	Johdanto	1
2	Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet	2
3	Henkilöstötutkimus	3
3.1	Henkilöstötutkimuksen sisältö ja toteutus	3
3.2	Henkilöstötutkimuksen tavoitteet	4
3.2.1	Parempi työntekijäkokemus	5
3.2.2	Parempi työhyvinvointi	6
3.2.3	Laadullisen tutkimuksen analysointi ja menetelmät	8
3.3	Valtion yhteinen henkilöstötutkimus.....	9
4	Tekoäly.....	11
4.1	Heikko ja vahva tekoäly	12
4.1.1	Heikko tekoäly	12
4.1.2	Vahva eli yleinen tekoäly.....	13
4.2	Data	13
4.3	Koneoppiminen	14
4.4	Syväoppiminen.....	15
4.5	Koneoppimisen menetelmiä	16
4.5.1	Ohjattu oppiminen	16
4.5.2	Vahvistusoppiminen.....	16
4.5.3	Ohjaamaton oppiminen	17
4.6	Luonnollisen kielen käsittely	17
4.6.1	NLU eli Luonnollisen kielen ymmärtäminen	19
4.6.2	NLP-tekniikat ja tekstin esikäsittely	19
5	Tekoäly käytännössä.....	21
5.1	Tekoälyn käytön vahvuudet	22
5.2	Tekoälyn käytön riskit	22
5.3	Tekoäly ja luonnollisen kielen käsittely käytännössä	23
5.4	Vastuullisuus tekoälyn käytössä	24
5.4.1	Yksityisyys.....	24
5.4.2	Vastuu.....	25
5.4.3	Vinouma	26
5.4.4	Datan laatu	26
6	Avovastausten analysoinnin selvitys	28

6.1	Selvityksen tavoite	28
6.2	Selvityksen toteutus ja tulokset	30
6.2.1	Asiakaskyselyn ja -haastattelujen suunnittelu ja toteutus	30
6.2.2	Asiakaskyselyn tulokset.....	32
6.2.3	Asiakashaastattelujen tulokset	35
6.2.4	Toimittajakyselyn suunnittelu, toteutus ja päätelmät.....	38
6.2.5	Nykyisen järjestelmän mahdollisuudet.....	39
6.3	Tekoälyn mahdollisuudet avovastausten käsittelyssä	40
7	Johtopäätökset ja pohdinta.....	43
8	Yhteenveto	45
	Lähteet.....	46

Kuvat ja taulukot

Kuva 1	Henkilöstötutkimuksen osa-alueet.....	3
Kuva 2	Kehitysympyrä	5
Kuva 3	Valtion yhteisen henkilöstötutkimuksen kysymysosiot	10
Kuva 4	Tekoälyn, koneoppimisen ja syväoppimisen suhde	11
Kuva 5	Yleisimmät vaiheet NLP mallin kouluttamisessa	20
Kuva 6	Henkilöstötutkimusten määrä vuodessa	32
Kuva 7	Tekoälyn tai automatisoinnin käyttö avovastausten käsittelyssä	34
Kuva 8	Avovastausten koosteet	34
Kuva 9	Koosteiden tekeminen itse vai ostaminen palveluna.....	35
Taulukko 1	Organisaation kokoon suhteutettu työmäärän keskiarvo	33

Liitteet

Liite 1	Aineistonhallintasuunnitelma
Liite 2	Asiakaskysely
Liite 3	Haastattelurunko

1 Johdanto

Henkilöstötutkimus antaa työnantajalle paljon tietoa henkilöstön mielipiteistä työstä, johtamisesta sekä työhyvinvoinnista. Kerätty tieto on tärkeää saada tehokkaasti työnantajan käyttöön. Yhtenä tietolähteenä käytetään henkilöstötutkimuksen avovastauksia, joiden hyödyntäminen on usein työlästä ja haastavaa. Laajoissa, vuosittain toteutettavissa tutkimuksissa avovastauksia voi kertyä paljon, niiden lukeminen vie paljon työaikaa ja lisäksi tulkitsemisessa voi korostua lukijan omat ajatukset.

Teknologia tarjoaa myös avovastauksien käsittelyyn tehokkaamman ratkaisun – tekoälyn. Siinä missä ihminen lukee vastauksia pitkiä aikoja, voi tekoäly käsitellä suurenkin aineiston nopeasti ja väsymättä. Mahdollisuuksia on monia, esimerkiksi vastausten jakaminen sentimenttien, teemojen tai yleisimpien sanojen mukaan.

Toimeksiantaja tuottaa asiakkailleen järjestelmän henkilöstötutkimusten toteuttamiselle sekä raportoinnille. Vuoden 2019 hallitusohjelman tavoitteena on ”Maailman paras julkinen hallinto”, jonka perusteella valtion yhteistä työtyytyväisyyskyselyä sekä kyselyn järjestelmää lähdettiin kehittämään (Valtiovarainministeriö ym., 2023, ss. 4, 14). Järjestelmän kehittämisen yhteydessä on noussut esiin tarve käsitellä avovastauksia tehokkaammin niin, että niitä voidaan hyödyntää paremmin henkilöstöjohtamisen kehittämisessä.

Opinnäytetyön tavoitteena on selvittää asiakkaiden tarpeita ja odotuksia avovastausten käsittelyyn sekä selvittää millaisia ratkaisuja nykyjärjestelmä avovastausten analysointiin tarjoaa. Lisäksi opinnäytetyössä selvitetään millaisia hyötyjä, mahdollisuuksia ja riskejä tekoälyn käyttöön liittyy.

Tutkimuskysymykset ovat:

- Mitä hyötyjä, mahdollisuuksia ja riskejä tekoälyn käyttöön sisältyy?
- Millaisia tarpeita asiakasorganisaatioilla on avovastausten analysointiin?
- Millaisia analysointimahdollisuuksia nykyinen ratkaisu tarjoaa avovastausten käsittelyyn?
- Millaisia analysointimahdollisuuksia markkinoilla olevat ratkaisut tarjoavat avovastausten käsittelyyn?

2 Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet

Opinnäytetyön toimeksiantaja Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet on valtionvarainministeriön alainen konsernipalveluiden tuottaja. Palkeiden ydintoimintaa on talous- ja henkilöstöhallinnon asiantuntijapalveluiden tuottaminen asiakkaille eli valtion hallinnon virastoille, rahastoille sekä laitoksille. Palkeet rahoittaa toimintansa omakustannusperiaatteella eli kustannukset katetaan palvelumaksuilla. (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-d)

Palkeiden palveluita tarjotaan asiakkaille yhtenäisillä järjestelmillä ja prosesseilla, ja niitä kehitetään yhteistyössä asiakkaiden kanssa. Tavoitteena on tarjota asiakkaiden arkea helpottavia henkilöstö- ja taloushallinnon palveluita. Henkilöstöpalveluihin kuuluu palkanlaskennan, palvelussuhteenhallinnan ja rekrytoinnin lisäksi osaamisen palvelut, joilla tuetaan asiakkaita valmentavassa työtöteessä ja työyhteisön kehittämisessä.

Taluspalveluissa asiakkaiden arkea tuetaan kirjanpidon, laskutuksen, tulotiliotteiden käsittelyn ja perinnän puolella. Lisäksi Palkeiden palveluvalikoimassa on useita erilaisia asiantuntijapalveluita, kuten tietojen analysointia ja ohjelmistorobotiikkaa. (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-e)

Valtion yhteinen henkilöstötutkimus on yksi Palkeiden tarjoamista palveluista. Valtion henkilöstötutkimus on vuosittain henkilöstölle toteutettava kysely, jolla kerätään henkilöstöltä palautetta johtamisen tueksi. Valtion yhteistä henkilöstötutkimusta ylläpidetään ja kehitetään Palkeissa. Henkilöstötutkimuspalvelun avulla asiakkaat voivat toteuttaa henkilöstötutkimuksia Palkeiden kanssa tai itsenäisesti, ja saavat käyttöönsä tutkimusmallit ja valmisraportit tutkimuksen tuloksista. Tutkimuksen tulokset tuottavat laajasti tietoa niin organisaatiotasolla kuin koko valtionhallinnosta. (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-a)

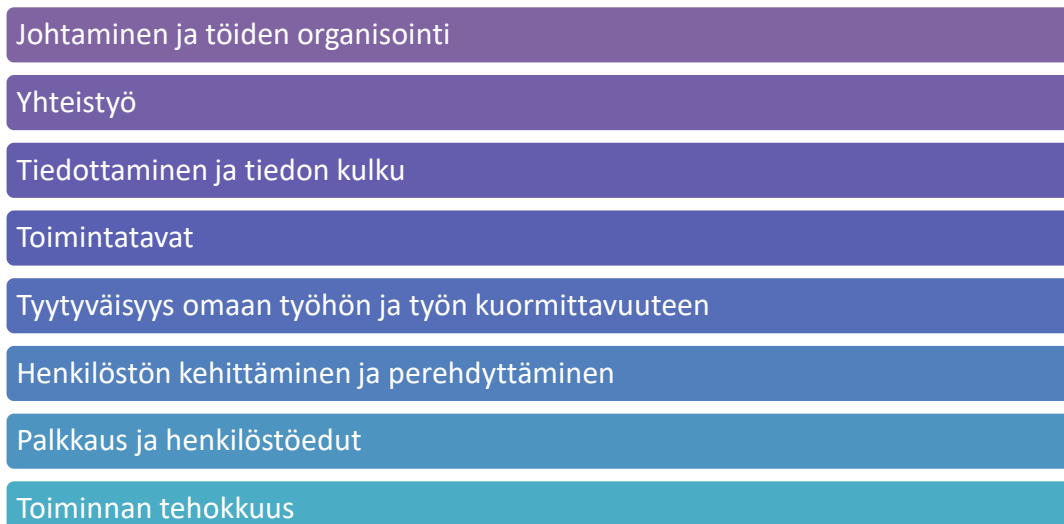
3 Henkilöstötutkimus

Työ ja työolosuhteet, esihenkilötyöskentely, tiedonkulku sekä työkyky ovat kaikki asioita, joita voidaan selvittää henkilöstötutkimuksella. Se onkin yleisnimitys tutkimuksille, joilla selvitetään henkilöstön mielipiteitä muun muassa työilmapiiristä, työhyvinvoinnista ja johtamisesta. Henkilöstötutkimuksen tavoitteena on kuvata henkilöstön työtyytyväisyyttä sekä löytää kehittämistarpeita. Säännöllisesti toteutettuna tutkimuksen tuloksista voidaan nähdä kuinka tehdyt kehittämistoimenpiteet ovat onnistuneet eri osa-alueilla. Toisaalta henkilöstötutkimuksella voidaan viestiä työntekijöille, että heitä kuunnellaan ja heistä välitetään. Tässä korostuu tutkimuksen tulosten jakaminen henkilöstölle ja organisaation toiminnan kehittäminen tulosten perusteella. (Joki, 2021, ss. 29–30)

3.1 Henkilöstötutkimuksen sisältö ja toteutus

Henkilöstötutkimusta voidaan kutsua myös työtyytyväisyystutkimukseksi tai ilmapiiritutkimukseksi, mikä kuvaa hyvin sen mahdollisuuksia ja tavoitteita. Usein henkilöstötutkimus tehdään säännöllisesti, esimerkiksi kerran vuodessa, jolloin siitä saadaan vertailutietoa. Vaikka henkilöstötutkimus usein painotetaan kohdeorganisaation mukaisesti, toistuvat tutkimuksissa Kuva 1 esittämät osa-alueet. (Hyppänen, 2013, kappale Henkilöstötutkimus, ensimmäinen ja toinen kappale)

Kuva 1 Henkilöstötutkimuksen osa-alueet (Hyppänen, 2013, Luku Henkilöstötutkimus, toinen kappale)



Henkilöstötutkimus on tyypillisesti määrällinen tutkimus, jota täydennetään laadullisilla kysymyksillä eli avoimella palautteella tai kehitysehdotuksilla (Huhta & Myllyntaus, 2021, s. 78). Tällaista tutkimusta kutsutaan monimenetelmätutkimukseksi (Juuti & Puusa, 2020, kappale 20 Monimenetelmätutkimus, ensimmäinen kappale). Tällöin tutkimuksen ensimmäinen osa antaa aiheen tutkimuksen toiselle osalle, mutta molemmilla tutkimuksen osilla on myös oma itsenäinen rooli tutkimuksen tuloksissa (Juuti & Puusa, 2020, kappale Tutkimusmenetelmien yhdistämistä koskeva käsitemaailma, neljäs kappale). Perinteisesti henkilöstötutkimusten määrällisessä osuudessa on käytetty väittämämuotoisia kysymyksiä, joihin vastataan Likertin asteikolla. Tällöin vastausvaihtoehdot ovat täysin eri mieltä, jokseenkin eri mieltä, ei eri eikä samaa mieltä, jokseenkin samaa mieltä ja täysin samaa mieltä. (Huhta & Myllyntaus, 2021, s. 82)

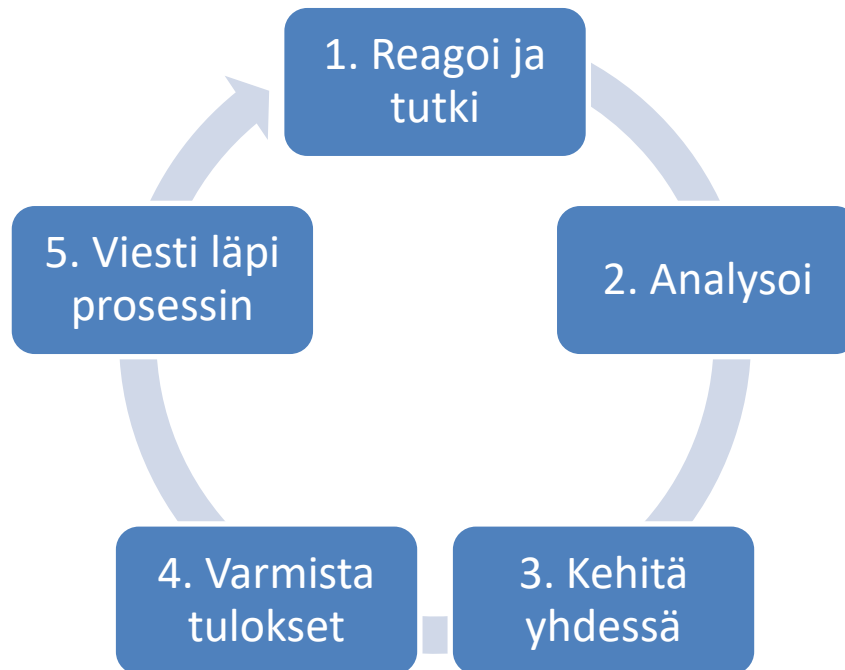
Koko organisaation kattavat henkilöstötutkimukset ovat vakiinnuttaneet paikkansa työelämässä jo 1970-luvulla ja niitä on pitkään toteutettu samantyyllisellä kehyksellä. Tällaisella kyselyllä saadaan kerättyä kattavasti tietoa koko organisaatiosta ja niiden mainostamiseen panostetaan. Tutkimuksen mainostamisella tutkimukselle usein saadaan korkea vastausprosentti ja tuloksia saadaan jopa tiimitasolle. Tämä tuo tutkimuksen tuloksia lähemmäs työntekijöitä ja mahdollistaa niin tiimin kuin esihenkilötyön kehittämisen. (Huhta & Myllyntaus, 2021, ss. 82–83)

3.2 Henkilöstötutkimuksen tavoitteet

Usein henkilöstötutkimuksen tavoitteena on rakentaa organisaation johdolle tilannekuvaa. On kuitenkin tärkeää huomata, että parempaan lopputulokseen pääsemisen kannalta tavoitteena tulee olla työntekijäkokemuksen kehittäminen. Tämä on huomioitava jo tutkimuksen suunnittelussa niin kysymysten kuin tulosten hyödyntämisen osalta. (Huhta & Myllyntaus, 2021, s. 83) Tutkimuksen suunnitteluvaiheessa on tärkeää miettiä tutkimusprosessia kokonaisuutena, kehityskohteiden tunnistamisesta tulosten viestintään. Huhta ja Myllyntaus (2021) ovat kuvanneet kehitysympyrässä (Kuva 2) vaiheet onnistuneelle tutkimukselle ja sen tuloksista tunnistetuille muutoksille. Kehitysympyrän vaiheita seuraamalla organisaatio voi jo ennen tutkimuksen aloittamista suunnitella miten tuloksia

halutaan analysoida ja esitellä henkilöstölle, miten tutkimuksen tuloksista nousseita asioita kehitetään ja miten tehtävistä muutoksista viestitään. (Huhta & Myllyntaus, 2021, ss. 79–80)

Kuva 2 Kehitysympyrä (Huhta & Myllyntaus, 2021, s. 80)



Työntekijäymmärryksen lisääminen, työntekijäkokemukseen panostaminen ja työhyvinvoinnista huolehtiminen maksaa itsensä takaisin. Parhaan asiakaskokemuksen luovat työntekijät, jotka voivat hyvin ja ovat innostuneita, motivoituneita sekä sitoutuneita. Lidl on kiteyttänyt tämän työnantajalupaukseensa: ”Onnelliset työntekijät takaavat myös onnellisimmat asiakkaat”. (Saramies & Törnroos, 2021, s. 41)

3.2.1 Parempi työntekijäkokemus

Huhta & Myllyntaus (2021, s. 38) määrittelevät työntekijäkokemuksen ajatuksiksi ja tunteiksi, joita työntekijässä herää vuorovaikutuksessa työnantajan kanssa.

Työntekijäkokemuksella kuvataan siis työntekijän ja työnantajan vuorovaikutuksessa muodostuneita tunteita ja ajatuksia työstä, työympäristöstä sekä työnantajaorganisaatiosta. Työntekijäkokemus voidaan nähdä sekä kokonaiskuvana eri vuorovaikutuspisteistä työnantajan kanssa että yksittäisinä tapahtumina, kuten esimerkiksi keskustelu esihenkilön kanssa. (Huhta & Myllyntaus, 2021, ss. 120–121) Työntekijän kokemuksiin, odotuksiin ja

havaintoihin vaikuttavat kaikki mitä hänen ympärillään tapahtuu, esimerkiksi organisaatiokulttuuri, organisaation arvot, johtaminen, työyhteisö ja työolot (Saramies & Törnroos, 2021, s. 39).

Henkilöstön sitouttaminen organisaatioon vaatii henkilöstön ymmärtämistä, tietoa siitä, miltä organisaatio työntekijöiden näkökulmasta näyttää ja mitkä asiat toimivat ja mitkä eivät. Vain henkilöstön ymmärtämisen kautta voidaan lähteä kehittämään työntekijäkokemusta. Ammattikirjallisuudessa henkilöstöä kuuntelevaa ja henkilöstön asemaan asettuvaa johtamistyyliä nimitetään palvelevaksi johtamiseksi. Palvelevan johtajan näkökulma on henkilöstön hyvinvoinnissa, potentiaalissa ja kasvussa. Ymmärtämällä ja kuuntelemalla henkilöstöä, organisaatioiden johto ja esihenkilöt voivat ymmärtää miten työntekijät toivovat heidät kohdattavan. Saramiehen ja Törnroosin mukaan (2021) tutkimuksessa on osoitettu HR-käytäntöjen ja johtamisen kehittämisen yhteys työtyytyväisyyteen: ”Kun yritys näyttää työntekijälle, että tämä on tärkeä, työntekijä vastaa huolenpitoon sitoutumalla ja olemalla tuotteliaampi”. (Saramies & Törnroos, 2021, ss. 39–40)

3.2.2 Parempi työhyvinvointi

Työntekijäkokemus vaikuttaa suoraan työhyvinvointiin. Mitä enemmän työntekijä kokee työnsä monipuolisena ja kokee pystyvänsä vaikuttamaan siihen, sitä paremmaksi hän kokee työhyvinvointinsa ja työn imun. (Saramies & Törnroos, 2021, s. 40) Organisaatioiden menestyksellä toiminta perustuu henkilöstöön, joka on innostunut, sitoutunut ja hyvinvoiva. Tähän tavoitteeseen pääseminen vaatii organisaatiolta ymmärrystä työntekijöiden toiveiden, tunteiden ja käyttäytymisen nykytilasta. Selvittämällä työhyvinvoinnin nykytilanne ja ennen kaikkea paneutumalla saatuihin tuloksiin, voidaan luoda organisaatiolle vahva selkäranka. (Saramies & Törnroos, 2021, s. 22)

Työhyvinvointi on laaja kokonaisuus, joka koostuu useasta eri osa-alueesta, kuten työ ja sen mielekkyys, terveys, hyvinvointi ja turvallisuus (Sosiaali- ja terveysministeriö, n.d.). Myös johtaminen, arvostus, palkitseminen, osaaminen, motivaatio sekä työntekijän voimavarat ja elämäntilanne vaikuttavat työhyvinvointiin (Hyppänen, 2013, Luku Työhyvinvoinnin kokonaisuus, ensimmäinen kappale). Työhyvinvointi on niin yksittäisen työntekijän kuin koko

työyhteisön kokemus ja sen edistäminen onkin työntekijöiden ja työnantajan yhteistyötä. Hyvä työhyvinvointi paitsi näkyy asiakkaille myös auttaa muun muassa tavoitteiden toteutumisessa, lisää joustavuutta ja uudistumiskykyä sekä parantaa työntekijäkokemusta sekä työyhteisön vetovoimaa. (Työturvallisuuskeskus, n.d.) Käytännössä työhyvinvointi vaikuttaa organisaatioissa kaikkeen. Kun työhyvinvointi on hyvällä tasolla, henkilöstön vaihtuvuus pienenee ja työntekijät ovat sitoutuneempia, työpaikan ilmapiiri on hyvä, asiakaskokemusten taso on parempi. Nämä kaikki vaikuttavat lopulta organisaation talouteen ja mahdollisuuksiin kasvaa ja kehittyä. (Great Place To Work, 2023)

Ilman tutkimustakin voidaan todeta, että hyvinvoiva ihminen saa aikaan enemmän ja parempia tuloksia kuin ihminen, joka voi huonosti. Se, parantaako työhyvinvointi liiketoiminnan tuloksia vai parantavatko hyvät liiketoiminnan tulokset työhyvinvointia, on tutkijoille epäselvää, koska liiketoiminnan lukuihin vaikuttavat myös monet muut muuttujat kuin työtyytyväisyys. Tästä huolimatta on voitu osoittaa, että menestyksekkäät organisaatiot voivat hyvin niin henkilöstötyytyväisyyttä kuin taloustilannetta katsoen. (Hyppänen, 2013, Luku Työhyvinvoinnin kokonaisuus, toinen kappale) Vuosina 2014–2017 Strategic Accounting -yrityksen, Turun ammattikorkeakoulun ja Vaasan yliopiston yhteistyössä tekemässä tutkimuksessa selvitettiin työilmapiirin vaikutusta yrityksen rahavirtoihin. Tutkimuksessa oli mukana 30 yritystä, jotka oli jaettu puoliksi hyvän ja huonon työilmapiirin kategorioihin työilmapiirikyselyjen perusteella. Tutkimuksessa voitiin osoittaa, että työilmapiirillä oli merkittävä vaikutus yrityksen taloudelliseen menestykseen. Vuonna 2015 taantuman taittuessa hyvän työilmapiirin yritysten rahavirtatulokset kasvoivat merkittävästi, kun heikon työilmapiirin yrityksissä ne puolestaan joko laskivat tai pysyivät samalla tasolla kuin ennenkin. Lisäksi vuonna 2017 hyvän työilmapiirin yritysten vuotuinen rahavirta oli noin 800 000 euroa suurempi kuin heikomman työilmapiirin yritysten. (Salonen ym., 2022, s. 142)

Huonolla tasolla olevasta työhyvinvoinnista aiheutuu huonon ilmapiirin, laadun ja tuloksellisuuden lisäksi konkreettisia kustannuksia. Sairauspoissaolot, työterveyshuollon sairauskäynnit, tapaturmat, varhainen eläköityminen tai osa-aikaeläke, vaihtuvuus sekä osaamisen poistuminen organisaatiosta aiheuttavat huomattavia suoria kustannuksia. Näiden lisäksi kustannuksia syntyy epäsuorasti esimerkiksi sijaisten hankkimisesta, uusien henkilöiden perehdyttämisestä sekä hiljaisen tiedon katoamisesta. Mikäli organisaatiossa on suurta vaihtuvuutta tai toistuvia sijaisjärjestelyjä, näkyy se myös ulospäin asiakkaille ja

sidosryhmille luotettavuuden pienenemisenä. Suuri työilmapiirin heikentäjä on työstressi ja työuupumus, jotka aiheuttavat poissaoloja, henkilöstön kynnistymistä sekä laadun ja luovuuden heikentymistä. Työilmapiiri kiristyy helposti, mikäli työntekijät ovat liian suuren kuormituksen alla, jolloin he eivät välttämättä jaksu tukea toisiaan tai olla hyväntuulisia. Huono työilmapiiri kumuloi lisää työyhteisöä ja henkilöstöä haittaavia tekijöitä, kuten epäasiallista kohtelua, häirintää tai henkistä väkivaltaa. Henkilöstönäkökulma on tärkeää ottaa huomioon niin yrityksen talouden kuin henkilöstön hyvinvoinnin takaamiseksi. (Joki, 2021, ss. 151–152)

3.2.3 Laadullisen tutkimuksen analysointi ja menetelmät

Laadullisen analysoinnin tavoitteena on saada tutkimusaineistosta enemmän tietoa ja tulkita mistä aineisto kertoo. Laadulliseen analysointiin on useita eri menetelmiä, mutta yhtä selkeää kaavaa analyysiin ei ole. Aineiston käsittelijän menetelmävalinnat vaikuttavat analyysin lopputulokseen ja analyysiä olisi pyrittävä tekemään pohjautuen valittuun menetelmään sekä teoretiseen tietoon. Perinteisiä sisällön analyysin menetelmiä ovat koodaaminen, teemoittelu ja tyypittely. (Günther ym., n.d.-a)

Juhilan mukaan (n.d.-a) koodaamisessa aineiston käsittelijä pyrkii saamaan aineiston helpommin hallittavaan muotoon yhdistelemällä ja erottelemalla aineiston osia eri ominaisuuksien mukaan. Teemoittelussa aineiston käsittelijä nostaa esiin tutkimuskysymysten kannalta merkittävät teemat ja usein toistuvat asiat. Teemoittelun tutkimusaineiston esittelyssä usein nostetaan esiin myös lainauksia aineistosta. (Juhila, n.d.-b) Tyypittelyssä aineiston käsittelijä puolestaan analysoi aineiston ilmiöitä ja pyrkii luokittelemaan niitä yhdistävien tekijöiden mukaan. Tähän voidaan hyödyntää esimerkiksi teemoittelua. Tyypit pyrkivät kuvaamaan aineiston keskeisimpiä ilmiöitä havainnollisesti. (Günther & Hasanen, n.d.-a)

Laadullisen analyysin ensivaiheina aineiston käsittelijä usein tutustuu aineistoon sekä muuntaa sen teknisesti käsiteltävään muotoon. Esimerkiksi ääninauhoitteet litteroidaan tekstiksi sekä kuva-aineistot skannataan yhteen tiedostoon. Tämän jälkeen aineistoa voidaan lähteä analysoimaan – lukemaan, jäsentelemään ja pohtimaan. Kokonaisuudessaan

laadullinen sisällön analyysi on monivaiheinen prosessi, jonka tavoitteena on löytää aineistosta uutta tietoa. (Günther ym., n.d.-a)

3.3 Valtion yhteinen henkilöstötutkimus

Valtion organisaatioilla on käytössä yhteinen henkilöstötutkimus, jolla kerätään tietoa muun muassa työtyytyväisyydestä, johtamistyöstä sekä palkkausjärjestelmän toimivuudesta. Henkilöstötutkimuksella kerätty tieto on keskeisessä roolissa tukemassa tiedolla johtamista. Oman organisaation tulosten lisäksi henkilöstötutkimuksesta tuotetaan vertailutietoa valtiotasoisesti ja hallinnonaloittain tukemaan organisaatioiden toiminnan parantamiseen. (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-ab)

Vuoden 2019 hallitusohjelman perustella käynnistettiin henkilöstöjohtamisen uudistusohjelma, jonka tavoitteena on parantaa valtion työnantajakuva työmarkkinoilla ja valmistautua suureen rekrytointitarpeeseen (Valtioneuvosto, 2019, s. 184). Valtion yhteistä henkilöstötutkimusta ja siihen käytettävää järjestelmää on lähdetty kehittämään uudistusohjelman myötä (Valtiovarainministeriö ym., 2023, s. 14) Vuoden 2023 aikana on kehitetty ja käyttöön otettu uutta ServiceNow-pohjaista henkilöstötutkimusjärjestelmää, jonka tavoitteena on vastata paremmin asiakkaiden tarpeisiin sekä uuteen palvelumalliin (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-f).

Valtion yhteinen henkilöstötutkimus koostuu kahdeksasta kysymysosioista (Kuva 3), joihin sisältyy yhteensä 43 yhteistä kysymystä sekä kysymysosiokohtaiset avokysymykset, joilla kerätään parannusehdotuksia kyseisen osion aiheisiin. Yhteisistä kysymysosioista lasketaan myös erilaisia indeksejä, esimerkiksi innovointikyvykkyyssindeksi, työyhteisöindeksi ja johtajuusindeksi. Organisaatioilla on mahdollisuus lisätä tutkimukselle myös omia kysymyksiä, jotka esitetään varsinaisten työtyytyväisyyskysymyksiäns jälkeen. Yhteisinä lisäkysymyksinä tutkimuksen lopussa kerätään tietoa halukkuudesta tehtäväkiertoon tai työpaikan vaihtoon, epäasiallisen käytöksen ja häirinnän kokemuksista, työhyvinvoinnista sekä työyhteisön vahvuuksista ja heikkouksista. (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-b)

Kuva 3 Valtion yhteisen henkilöstötutkimuksen kysymysosiot (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-b)

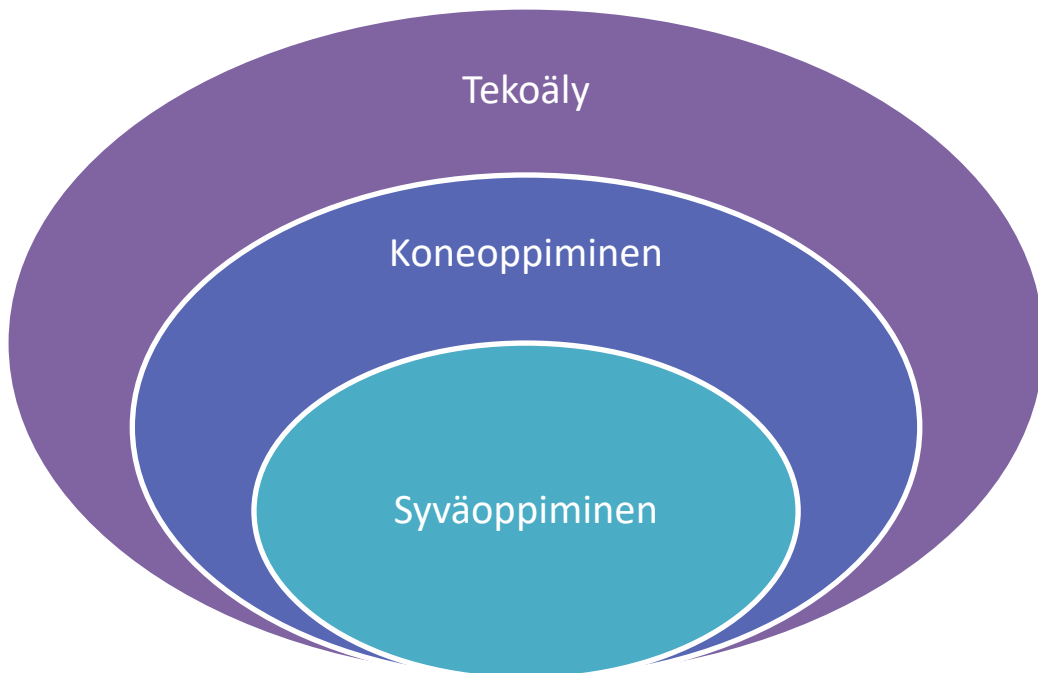


Valtion yhteiselle henkilöstötutkimukselle on tarjottu erilaisia raportointimalleja, joita voidaan käyttää tutkimuksen päätyttyä. Näitä raportointimalleja ovat vastausjakaumat, keskiarvot sekä vapaat tekstivastaukset. Raporteille voidaan lisätä vertailutiedoksi esimerkiksi koko organisaation tai valtion tulokset. (Valtion talous- ja henkilöstöhallinnon palvelukeskus, n.d.-c)

4 Tekoäly

Tekoäly, AI, Artificial Intelligence, keinoäly, koneoppiminen ja syväoppiminen ovat kaikki samaan aiheeseen liittyviä termejä ja puhekielessä ne sekoittuvat helposti keskenään (Kananen & Puolitaival, 2019, s. 27). Neljä ensimmäistä termiä ovat keskenään synonyymejä, joilla tarkoitetaan ihmisen toimintaa ja päätöksentekoa jäljittelevää ohjelmistoa. Hyvä esimerkki tekoälyn käyttötarkoituksesta on suurien datamäärien käsittely, mihin ihminen ei pystyisi tai siihen menisi vuosia aikaa. Koneoppimisen sekä syväoppimisen keinoilla tekoäly voidaan opettaa oppimaan uutta ja parantamaan suorituskyykyään ainakin osittain itsenäisesti. Koneoppimisessa tekoäly käyttää hyväkseen dataa, jonka perusteella se oppii uusia asioita joko itsenäisesti, ohjatusti tai vahvistetusti eli ihmisen antamalla palautteella tekoälyn tekemistä päätöksistä. Syväoppiminen on yksi koneoppimisen muodoista, jossa hyödynnetään neuroverkkoja asioiden sekä kontekstien tunnistamiseen. (Kolari & Kallio, 2023, Luku Perustermit tutuiksi, kappaleet 2–3, 7, 8) Klemettiä (2022) mukaillen Kuva 4 on havainnollistettu tekoälyn, koneoppimisen sekä syväoppimisen suhdetta toisiinsa.

Kuva 4 Tekoälyn, koneoppimisen ja syväoppimisen suhde (Klemetti, 2022)



Tekoäly koostuu ohjelmoinnin, matematiikan ja tilastotieteen useista menetelmistä ja tekniikoista, joista pyritään valitsemaan tehokkain väline ratkaistavaan ongelmaan.

Perusteet tekoälyn toiminnassa ovat periaatteen tasolla helposti ymmärrettäviä ja tuttuja matemaattisia käsitteitä. Käytäntö on kuitenkin monimutkaisempi, koska ihmisen on vaikeaa ymmärtää kolmiulotteista avaruutta moniulotteisempia avaruuksia, kun taas tekoälylle on nopeaa ja tehokasta käsitellä moniulotteisia ja -tasoisia muuttujia rajattomassa määrässä ulottuvuuksia. (Kananen & Puolitaival, 2019, s. 27)

4.1 Heikko ja vahva tekoäly

Termit tekoäly tai Artificial Intelligence tuovat mieleen koneen, joka osaa ajatella. Tämä ei kuitenkaan pidä paikkaansa, sillä tämänhetkisellä tekoälyllä ei ole älykkyyttä samassa merkityksessä kuin ihmisillä ajatellaan olevan. (Chishti ym., 2020, ss. 39–40) Äly on epämääräinen käsite eikä sitä voida mitata kiistattomasti, kuten esimerkiksi nopeutta, mikä vaikeuttaa tekoälyn määrittelemistä. Voidaan kuitenkin ajatella, että älykkyys tarkoittaa uuden oppimista sekä ongelmanratkaisukykyä opittuja taitoja soveltaen. (Järvinen, 2023, s. 49) Yleisesti ottaen tekoäly voidaan jakaa vahvaan ja heikkoon tekoälyyn. Kaikki tällä hetkellä käytössä olevat tekoälyt ovat heikkoa tekoälyä. Vahvaa tekoälyä ei ole ainakaan vielä pystytty kehittämään. (Kananen & Puolitaival, 2019, s. 38)

4.1.1 Heikko tekoäly

Kun puhutaan tekoälystä, viitataan sillä kapeaan eli heikkoon tekoälyyn, sillä yleistä tekoälyä ei ole vielä kehitetty. Termillä kapea, kuvataan sitä, että tekoälyä voidaan käyttää yksittäisessä määritellyssä tehtävässä, mutta se ei voi laajentaa osaamistaan muihin tehtäviin. Esimerkiksi kasvilajeja luokitteleva kuvantunnistustekoäly ei pysty vastaamaan kysymyksiin asiakaspalvelussa. (Merilehto, 2018, ss. 23–24) Heikkoa tekoälyä voidaan käyttää siis selkeissä ja tarkasti rajatuissa tehtävissä ja sen vahvuuksiin kuuluu esimerkiksi datamassojen käsittely, toistoa sisältävät tehtävät sekä asioiden järjestely. Tällaisissa tehtävissä se pystyy suoriutumaan jopa ihmistä paremmin ja nopeammin. (Kananen & Puolitaival, 2019, ss. 37–38)

4.1.2 Vahva eli yleinen tekoäly

Vahva tekoäly osaa ajatella ja toimia itsenäisesti ja sillä on inhimillinen ymmärrys (Kolari & Kallio, 2023, Luku Perustermit tutuiksi, kappale 3). Vahvalla tekoälyllä tarkoitetaan yleiskäyttöistä tekoälyä, jonka toimintaa voidaan verrata ihmiseen. Se osaa soveltaa taustatietoja ja tehdä ihmiselle ominaisia asioita. Usein ihmisten tekoälyyn liittyvät pelot koskevat vahvaa tekoälyä, josta on saatu mediassa suureellisia otsikoita. (Kananen & Puolitaival, 2019, s. 38) On hyvä muistaa, että vahvaa tekoälyä ei ole vielä onnistuttu kehittämään eikä ole takeita, että siihen päästään koskaan (Coeckelbergh, 2020/2021, s. 74).

Vaikka koneiden on helppo löytää säännönmukaisuuksia tai malleja datasta sekä luoda ennusteita dataan perustuen, on niille erittäin vaikeaa, ellei mahdotonta, ymmärtää asiayhteyksiä ja kokonaisuuksia. Jotta vahvaa tekoälyä pystyttäisiin kehittämään, tulisi tekoälyn pystyä itse vastaamaan kysymykseen miksi. Eli käytännössä syyn ja seurauksen suhde pitäisi pystyä kuvaamaan yksiselitteisesti. Tällä hetkellä tekoälyn toiminta pohjautuu suureen datamassaan, minkä perusteella se pystyy vastaamaan kysymyksiin mitä ja miten, mutta ei miksi. (Kananen & Puolitaival, 2019, ss. 40–41) Toisaalta tekoälyllä pitäisi olla inhimillistä kokemusta ja usein myös erehdyksistä oppimista, jotta sen olisi mahdollista ymmärtää ympäröivää maailmaa. Järvinen käyttää esimerkkinä hissien hälytysnappia, jonka vieressä yleensä lukee: ”paina nappia vähintään 5 sekuntia”. Ihmiset ymmärtävät, että ohje koskee hätätilanteita mutta tekoälyllä varustettu robotti painaisi nappia kyseenalaistamatta ohjetta. Yhdyskunta ei kuitenkaan salli koneille virheiden tekemistä eikä näin ollen oppimista erehdyksien kautta. (Järvinen, 2023, s. 51)

4.2 Data

Keskeisessä roolissa tekoälyn käyttöön otossa ja kouluttamisessa on data. Data voi olla useissa eri muodoissa, kuten tekstinä, numeroina tai kuvina. Oleellista on, että data on laadukasta eli yksiselitteistä ja yhdenmukaista. Tekoäly voi käsitellä vain laadukasta dataa. Kun lähdetään määrittelemään ja käyttöönottamaan tekoälyjärjestelmää, on tärkeää huomioida, millaista dataa tarvitaan, jotta sitä voidaan hyödyntää analysoinnissa. Myös järjestelmän käyttäjien yhdenmukainen ohjeistaminen järjestelmän käyttöön edesauttaa laadukkaan datan muodostamista. (Kananen & Puolitaival, 2019, ss. 71–72)

Ennen kuin tekoäly voi hyödyntää dataa, se on kerättävä ja puhdistettava sekä tunnistettava mikä data on olennaista eli mitä aineistoa käytetään. On myös esimerkiksi mietittävä, miten dataa eli tietoaimeistoa saadaan kerättyä riittävästi, miten se yhdistetään ja miten se rakennetaan uudelleen. Kaikkien näiden dataan liittyvien ongelmien ratkaiseminen on osa datatieteen kenttää, joka kuuluu vahvasti koneoppimiseen, jolla suuria data-aineistoja voidaan analysoida automaattisesti. Niin datatiede kuin koneoppiminen pohjautuvat tilastotieteeseen, jossa puolestaan on tavoitteena yhdistää aineiston yksittäiset havainnot yleiseksi kuvaukseksi. (Coeckelbergh, 2020/2021, s. 94)

4.3 Koneoppiminen

Koneoppiminen on yksi tekoälyn osa-alue. Se poikkeaa perinteisestä ohjelmoinnista, jossa sovellukselle annetaan sääntöjä, joiden perusteella sovellus tuottaa lopputuloksen. Koneoppimisessa sovellukselle ei anneta sääntöjä vaan dataa. (Merilehto, 2018, ss. 27–28)

Nimitys koneoppiminen on ristiriitainen, koska voidaan ajatella, etteivät koneet voi todella oppia. Koneoppiminen ei siis ole samanlaista oppimista kuin ihmisten oppiminen. Se on tilastotieteellinen prosessi ja perustuu algoritmeihin, joille annetaan aineisto, josta se pyrkii tunnistamaan sääntöjä ja hahmoja. Näillä tunnistetuilla säännöillä sekä hahmoilla algoritmit selittävät aineistoa sekä ennustavat niiden perusteella tulevia aineistoja. (Coeckelbergh, 2020/2021, s. 89) Nämä algoritmit tarvitsevat paljon dataa eikä niiden suorituskyky tai tulokset parane suhteessa datan määrään (Kananen & Puolitaival, 2019, s. 109). Algoritmi puolestaan on ohje ongelman ratkaisemiseksi. Se ilmaisee, kuinka asia tulee tehdä halutun lopputuloksen saavuttamiseksi. Algoritmeja on useita erilaisia ja niitä käytetään eri käyttötarkoituksiin. Ei siis ole yhtä ja ainoa algoritmia, joka toimisi kaikkialla, vaan useimmiten käytännössä hyödynnetään useampia algoritmeja yhden ongelman ratkaisemisessa. (Kananen & Puolitaival, 2019, s. 112)

Käytännössä koneoppiminen voi tarjota ratkaisuja moniin liiketoiminnan ongelmiin ja sitä voidaan hyödyntää erilaisissa sovelluksissa. Jos koneoppimista hyödyntävälle sovellukselle annetaan syötteenä esimerkiksi valokuva, se voi antaa vasteena kuvatekstin valokuvalla, mitä voidaan hyödyntää kuvien merkitsemisessä. Toisena esimerkkinä sovellukselle voitaisiin antaa ääninauhoite, jonka perusteella se litteroi tekstin eli tunnistaa puheen. (Merilehto,

2018, s. 29) Riittävälle tasolle koulutettuihin koneoppimisen malleihin voidaan automaation avulla yhdistää myös hälytyksiä, jotka ilmoittavat käyttäjälle poikkeamista tai havaituista muutoksista. Esimerkiksi malli voi ilmoittaa asiakastyytyvyyden laskusta lähes reaaliajassa, jolloin siihen voidaan reagoida nopeasti. (Merilehto, 2018, s. 33)

4.4 Syväoppiminen

Yksi koneoppimisen osa-alue on syväoppiminen, joka perustuu neuroverkkoihin (Carpesato, 2020, s. 19) Syväoppimisen voidaan ajatella olevan neuroverkkojen uusi tuleminen, koska neuroverkot olivat suosittuja jo 1980- ja 1990-luvuilla (Merilehto, 2018, s. 56). Toimiakseen neuroverkot tarvitsevat paljon dataa sekä laskentatehoa, minkä vuoksi neuroverkot ovat yleistyneet vasta 2010-luvulla, jolloin niitä on ollut riittävästi saatavilla sekä osaaminen matemaattisten ideoiden käytäntöön toteuttamiseen on lisääntynyt. Neuroverkot ovat monimutkaisia kokonaisuuksia ja niiden rakenne on biologisesti samankaltainen kuin ihmisen aivoilla, mutta kuitenkin neuroverkot eivät ole samanlaiset kuin ihmisen aivot. Neuroverkkojen tekninen ymmärtäminen ei ole välttämätöntä, jotta tekoälyä voidaan hyödyntää liiketoiminnassa, mutta perusteiden hallitseminen auttaa löytämään kohteita, joissa tekoälyä on mahdollista hyödyntää. (Kananen & Puolitaival, 2019, ss. 127–128)

Neuroverkot koostuvat yksittäisistä neuroneista ja niiden välisistä yhteyksistä. Yksittäisille neuroneille on annettu matemaattisen funktion kaava, joka saa herätteen (input) ja se palauttaa lähtöarvon (output). Yksittäiset neuronit on kytketty toisiinsa ja jokaisessa kytkennässä on painokerroin, joka kertoo kuinka paljon neuronin antama lähtöarvo vaikuttaa tutkittavaan ilmiöön. Käytännössä painokertoimet siis ilmaisevat kuinka vahva kahden asian yhteys on ja tähän perustuu koko neuroverkon niin kutsuttu äly. Kun neuroverkkoa koulutetaan, etsitään juuri oikeita painokertoimia, joita on lukemattomia. Koulutus algoritmi tunnistaa datasta sääntöjä, jotka ovat neuroverkon painokertoimia. (Kananen & Puolitaival, 2019, ss. 129–133)

Neuroverkkoja on erityyppisiä, joita käytetään eri tarkoituksiin. CNN (convolutional neural network) eli konvoluutivneuroverkkoa hyödynnetään kuvien käsittelyyn, vastaavasti kuin ihminen ottaa vastaan tietoa kuvista. (Nelson, 2019, s. 9) Toinen yleinen syväoppimisen muoto on RNN eli recurrent neural network, suomennettuna takaisinkytketty neuroverkko.

Se perustuu koulutusdataan, jossa aineiston esimerkit ovat merkitty, luokiteltu tai jaettu kategorioihin. (Nelson, 2019, s. 17)

4.5 Koneoppimisen menetelmiä

Koneoppiminen voi olla ohjattua oppimista, vahvistusoppimista tai ohjaamatonta oppimista. Kaikissa kolmessa vaihtoehdossa ihminen on jollakin tavalla mukana ja jokaisessa vaihtoehdossa myös algoritmi toimii jossakin määrin itsenäisesti. (Coeckelbergh, 2020/2021, ss. 91–92) Data on suuressa roolissa koneoppimisessa. Ohjatussa oppimisessa on paljon merkittävää dataa ja vahvistusoppimisessa kokeillaan, annetaan palautetta ja sen jälkeen parannetaan toimintaa. Ohjaamattomassa oppimisessa taas on paljon dataa, joka klusteroidaan eli jaetaan ryhmiin, joissa on mahdollisimman paljon samankaltaisia tietoja (Itä-Suomen yliopisto, 2020). (Campesato, 2020, s. 19)

4.5.1 Ohjattu oppiminen

Koneoppimisessa hyödynnetään usein ohjattua oppimista, jossa koneoppimisen mallia opetetaan valmiilla opetusdatalla. Opetusdataan on merkitty valmiiksi syöte (input) ja vaste (output), joiden perusteella malli oppii ennustamaan oikeaa lopputulosta. Tämän jälkeen mallia voidaan vielä testata opetusdataa vastaavalla testidatalla. Näin voidaan selvittää, onko mallin koulutus onnistunut. (Merilehto, 2018, ss. 28–29) Ohjatussa oppimisessa on tärkeää, että koulutusdatassa on sekä oikeita että vääriä esimerkkejä. Käytännössä dataa tulee olla tilanteista, jolloin kaikki menee suunnitellusti sekä erilaisista poikkeustilanteista, joissa syystä tai toisesta kaikki ei mene suunnitelman mukaan. (Merilehto, 2018, s. 56)

4.5.2 Vahvistusoppiminen

Myös vahvistusoppiminen on koneoppimisen osa-alue, jossa tekoälysovellus oppii yrityksen ja erehdyksen kautta. Vahvistusoppimisessa tekoälysovellus saa positiivista tai negatiivista palautetta eli niin sanotusti palkinnon tai rangaistuksen. Tekoäly oppii virheistään ja pyrkii tavoittelemaan mahdollisimman hyvää palkintoa. Jos vahvistusoppimista verrataan ohjattuun oppimiseen, on niillä yhteistä ainoastaan syöte ja vaste. Ohjatussa oppimisessa

palautteessa annetaan oikea vastaus, mutta vahvistusoppimisessa palautteessa ei ole oikeaa vastausta, vaan sovellus itse tekee päätöksen, kuinka korjata toimintaansa. (Brooks, 2021)

4.5.3 Ohjaamaton oppiminen

Ohjaamattoman oppimisen menetelmiä on käytetty menestyksekkäästi esimerkiksi kielen kääntämisessä tai kasvojen ja kuvien tunnistuksessa. Ohjaamattomassa oppimisessa ohjelmalle annetaan esimerkiksi kuvia, joista se tulkitsee havaintoyksiköiden ominaisuuksia ja niiden perusteella tekee ryhmittelyn, kokonaan ilman vertailua ennalta annettuihin lopputuloksiin. (Raatikainen, 2021, s. 120) Ohjaamattomassa oppimisessä ihminen ei siis valmiiksi kerro tekoälylle haluttuja lopputuloksia, joihin esimerkiksi kuvat luokiteltaisiin, vaan tekoäly muodostaa itse saamastaan datasta luokat, joihin annettu data jaetaan. Tekoälyn muodostamat luokat voivat olla epäloogisia tai sitten ne ovat järkeviä ja niiden avulla saadaan täysin uutta tietoa tekoälylle annetusta aineistosta. (Coeckelbergh, 2020/2021, ss. 91–92)

4.6 Luonnollisen kielen käsittely

Ihmisen kieltä ei ole tarkoitettu koneiden luettavaksi, mistä on syntynyt tarve luonnollisen kielen käsittelyn menetelmille. Luonnollisen kielen käsittely on kehittynyt lähiaikoina paljon, kun päättelyyn perustuvia menetelmiä on pystytty korvaamaan koneoppimisen menetelmillä. Monelle tutut Open AI:n GPT-3 tai Googlen BERT ja LaMDA ovat jo kehittyneet niin pitkälle, että ne voivat tulkita konteksteja sekä keskustella uskottavasti ihmisen kanssa. (Kolari & Kallio, 2023, kappale Tekoälyn osa-alueet, Luonnollisen kielen käsittely)

Luonnollisen kielen käsittelystä käytetään yleisesti termiä NLP, joka tulee englannin kielen sanoista Natural Language Processing. Luonnollisen kielen käsittelyllä eli NLP:llä viitataan niin kirjoitetun kuin puhutun kielen käsittelyyn. Käsittelyllä tarkoitetaan luokittelua, generointia tai keskustelua. (Kananen & Puolitaival, 2019, s. 141) Jotta tekstiä voidaan hyödyntää, se täytyy siis analysoida ja jäsentää. NLP tekee annetuille tiedoille kielellisen analyysin usealla eri tasolla. Näillä tasoilla se esimerkiksi järjestelee sanat ja lauseet niin, että se voi ymmärtää ja vertailla niitä, selvittää lauseiden syntaksin ja semantiikan, tunnistaa

sanat eri taivutusmuodoissa ja lopulta tehdä erilaisia tulkintoja muun muassa sanojen ja lauseiden sentimenteistä. (Chishti ym., 2020, ss. 46–47)

NLP on saanut alkunsa jo 1900-luvun puolivälissä ja kehittynyt sääntöpohjaisista sovelluksista koneoppimiseen ja syväoppimiseen. Ensimmäinen kehitysvaihe kesti aina 1990-luvulle saakka ja perustui suurelta osin ehdolliseen logiikkaan, minkä haasteena olivat esimerkiksi erilaiset lauserakenteet, prepositiot, adverbit sekä lauseiden välinen yhteys. Ensimmäisestä vaiheesta on päästy tilastollisen analyysin ja koneoppimisen menetelmien kautta nykyiseen kehitysvaiheeseen. Käänteentekevä vuosi oli 2012, jolloin konvoluutioneuroverkot saavuttivat läpimurron kuvien luokittelun tarkkuudessa, minkä seurauksena tutkijat löysivät tapoja hyödyntää niitä NLP:n tehtävissä. Konvoluutioneuroverkot kehittivät sittemmin syväoppimisen menetelmiksi: takaisinkytketyiksi neuroverkoiksi ja pitkän lyhytaikaisen muistinkäytöksi. Vuoden 2017 lopulla myös nämä syväoppimisen menetelmät kehittivät transformer-arkkitehtuureiksi, jotka yltyivät aiemmin näkemättömään suorituskyykyyn luonnollisen kielen käsittelyn kentällä. (Campesato, 2021, ss. 54–55)

Jotta kone pystyy käsittelemään myös poikkeustilanteita ja erilaisia kielen sävyjä, tulee sen ymmärtää myös kielioppisääntöjä. Näitä tilastotieteessäkin tuttuja haasteita on lähdetty käsittelemään luokitteluongelman näkökulmasta. Kananen ja Puolitaival esittävät esimerkin, jossa koneelta kysytään esimerkiksi, onko annettu substantiivi monikko tai esiintyykö sana tekstissä ensimmäistä kertaa, joihin kone antaa vastauksena joko kyllä eli 1 tai ei eli 0. Vastausten perusteella teksti voidaan muuttaa vektoreiksi [0, 1, 1, 0...], jolloin lauseiden voidaan ajatella olevan havaintoja pisteavaruudessa. Kone siis muuttaa sille annetun tekstin vektoreiksi, laskee sanoille sijainnin ja sanojen esiintymien määrän, minkä perusteella se pystyy antamaan halutun vastauksen. Sanojen muuttaminen vektorimuotoon on parantanut tekoälyn kykyä tunnistaa samankaltaisia sanoja, kun se laskee vektorien välisiä etäisyyksiä. Myös kielen kääntäminen perustuu vektoreihin, kun kone muuttaa annetun tekstin vektoreiksi, laskee etäisyydet ja muuntaa vektorit sanoiksi kohdekielellä. (Kananen & Puolitaival, 2019, ss. 143–144)

4.6.1 NLU eli Luonnollisen kielen ymmärtäminen

Luonnollisen kielen käsittelystä voidaan tunnistaa kaksi eri alaluokkaa; luonnollisen kielen ymmärtäminen eli NLU Natural Language Understanding ja luonnollisen kielen tuottaminen eli NLG Natural Language Generation. NLU mahdollistaa koneille luonnollisen kielen ymmärtämisen ja analysoinnin poimimalla esimerkiksi käsitteitä, kokonaisuuksia, avainsanoja tai tunteita. Siihen kuuluu monia kielitieteen termejä ja NLP:n tasoja, esimerkiksi syntaksi eli lauseoppi, semantiikka eli merkitysoppi ja pragmatiikka, joka keskittyy yleisen tiedon hyödyntämiseen. (Khurana ym., 2023, ss. 2–6) Yksi NLU:n tärkeistä tehtävistä on erottaa merkitysten yhteydet tekstistä, jotka voivat olla peräisin eri lähteistä kuten dokumenteista tai chatbotilta. Tekstissä voi esimerkiksi olla sana 'hän' useaan kertaan, jolloin NLP:n tehtävä on tulkita, viitataanko sanoilla samaan vai useampaan eri henkilöön. Tällaiset ja vastaavat merkitysten moniselitteisyydet voidaan luokitella eri tyyppeihin, kuten leksikaaliseen tai syntaksiseen. (Carpesato, 2021, ss. 57–58) Leksikaalisessa moniselitteisyydessä yksi esimerkki tekstin käsittelytavasta on puheosatunnisteen lisääminen jokaiselle sanalle. Tässä käsittelytavassa NLP pyrkii tunnistamaan sanan kontekstin silloinkin, kun sanalla on monta merkitystä ja lisäämään sanalle tunnisteiden. Syntaksisessa moniselitteisyydessä on puolestaan kyse lauserakenteesta ja sanojen välisistä yhteyksistä. Syntaksin osalta käsittelyssä pyritään tunnistamaan vaikuttaako sanajärjestyksen muuttaminen lauseen merkitykseen. (Khurana ym., 2023, ss. 4–5)

4.6.2 NLP-tekniikat ja tekstin esikäsitteily

Luonnollisen kielen käsittelyllä on monia erilaisia tehtäviä, joten on myös useita tekniikoita, joita voidaan hyödyntää erilaisiin käyttötarkoituksiin. (Carpesato, 2021, s. 61) Yksi tekniikoista on ohjaamattoman koneoppimisen menetelmä aihehallinnus, joka perustuu dokumenteissa esiintyvien sanojen esiintymistiheyteen. Tässä tekstidokumenttien ryhmittelymenetelmässä pyritään tunnistamaan dokumentit, jotka ovat samankaltaisia keskenään. Menetelmässä aiheita ei määritellä ennalta, mutta aiheiden määrä määritellään. Lopputuloksena saadaan listaus jokaisen aiheen yleisimmistä sanoista sekä tieto dokumenttien kuulumisista kyseisiin aiheisiin. (Ropponen, 2022, ss. 7–8)

Tekniikoiden lisäksi on myös erilaisia vaiheita NLP mallin kouluttamiseen. Myös nämä vaiheet riippuvat tehtävästä, joka mallin tulee suorittaa. Kuva 5 nähdään yleisimmät vaiheet, joita NLP mallien kouluttamisessa käytetään. (Campesato, 2021, ss. 61–62) Tekstiaineiston esikäsittelyyn kuuluu tarpeettomien sanojen ja merkkien poistaminen. Tällaisia ovat esimerkiksi hukkas sanat (eli muun muassa mutta, ei, olla sekä tai), välimerkit, numerot, sähköposti- ja internetosoitteet sekä erisnimet. (Ropponen, 2022, s. 4) Voidaan katsoa, että kirjainten muuntaminen pieniksi kirjaimiksi ja melun poisto eli tekstinkäsittelyä häiritsevien merkkien, numeroiden sekä sanojen poisto, ovat pakollisia vaiheita. Tekstin normalisointi eli esimerkiksi erikoismerkkien, emoji- ja hyperlinkkien poisto on erittäin suositeltava vaihe. Tekstin rikastaminen, hukkasanojen poisto, karsinta ja perusmuotoistaminen puolestaan ovat riippuvaisia suoritettavasta tehtävästä. (Campesato, 2021, ss. 61–62)

Kuva 5 Yleisimmät vaiheet NLP mallin kouluttamisessa (Campesato, 2021, s. 61)

Sanojen muuntaminen pieniksi kirjaimiksi

Melun poisto

Normalisointi

Tekstin rikastaminen

Hukkasanojen poisto

Stemming = karsinta

Lemmatization = perusmuotoistaminen

5 Tekoäly käytännössä

Tekoälyllä voidaan parantaa organisaation nykyisiä prosesseja, luoda uusia palveluita olemassa olevasta datasta tai käyttää dataa tiedolla johtamiseen ja päätöksenteon tueksi. Tämä edellyttää, että tunnetaan organisaation prosessit, asiakkaat ja markkinat. Lisäksi tekoälyä käyttöönotettaessa on kannattavaa huomioida sen tuomat tuotot ja investoinnin takaisinmaksuaika eli tehdä arvoanalyysi. Kun organisaatiossa pohditaan tekoälyn mahdollisia käyttökohteita, tarvitaan organisaation toiminnan asiantuntijoiden sekä operatiivisen johdon yhteistyötä. (Kananen & Puolitaival, 2019, ss. 199–200)

Tekoälyn käyttökohteeksi voidaan tunnistaa organisaation sisäinen prosessi, jota halutaan tehostaa. Jos prosessissa on paljon toistoa ja säännönmukaisuutta, se sopii todennäköisesti hyvin tekoälyn käyttökohteeksi. Kun tekoälyä hyödynnetään sisäisiin prosesseihin, on tarpeellista huomioida sen vaikutus koko organisaation tasolla, koska usein tekoäly nopeuttaa prosessiin kuluva aika. Kun yhteen prosessiin kuluva aika lyhenee, miten se vaikuttaa muihin prosesseihin? Onko tarpeellista tehostaa myös muita prosesseja tai organisoida niitä toisin? Ovatko organisaation toimintamallit ketteriä ja tekoälyn tuomat hyödyt saadaan todellisesti käyttöön? Mikäli koko organisaation prosesseja ei ole tiedostettu tekoälyn käyttöönotossa, voi yhden prosessin tehostaminen aiheuttaa ruuhkaa seuraavaan prosessiin mikä puolestaan voi pidentää kokonaisuudessaan tehtäviin käytettyä aikaa. Siksi huolellinen prosessien läpikäynti on tärkeää, jotta tekoälyllä tavoiteltavat hyödyt toteutuvat. (Kananen & Puolitaival, 2019, ss. 200–201)

Tekoälyä on myös mahdollista käyttää kokonaan uuden palvelun tai tuotteen kehittämiseen, kuten myös olemassa olevien palveluiden ja tuotteiden parantamiseen. Kun suunnitellaan uutta tekoälypohjaista tuotetta tai palvelua, on kiinnitettävä erityistä huomiota siihen missä laajuudessa tehtäviä annetaan tekoälylle. Vaativiakin asiantuntijatehtäviä voidaan kouluttaa tekoälylle, mutta joskus tekoälyn tekemät tulokset vaativat ihmisen hyväksynnän. (Kananen & Puolitaival, 2019, s. 202)

5.1 Tekoälyn käytön vahvuudet

Kun tekoäly on käytössä selkeästi rajatuissa tehtävissä, se on nopeampi ja tarkempi kuin ihminen vastaavassa tehtävässä. Se ei myöskään väsy eli voi tehdä loputtomasti toistoja ja on asioista samaa mieltä ajasta riippumatta. Etuna tekoälylle voidaan katsoa myös puolueettomuus, kun huomioidaan, että myös sen koulutukseen käytetyn datan tulee tällöin olla puolueetonta. (Kananen & Puolitaival, 2019, s. 37) Tekoäly voi työskennellä vuorokauden ympäri viikon jokaisena päivänä. Tekoäly ei tee inhimillisiä virheitä. Tekoälyn suoritus on tasavarma. Tasalaatuisuus, toistuvuus ja edulliset kustannukset riittävät joissakin tehtävissä ihmisen päihittämiseen. (Järvinen, 2023, s. 185)

Ihmiselle vaikeinta työssä on aloittaminen. Saatavilla on sosiaalisen median keskusteluita, videoklippejä ja verkkokauppoja, jotka kaikki häiritsevät keskittymistä. Tekoälyllä tällaista ongelmaa ei ole. Keskittymisen lisäksi tekoäly on ylivoimainen moniajossa eli multitaskingissa. Tietokoneiden prosessorien sisällä on useita ytimiä, jotka todellakin tekevät eri tehtäviä yhtä aikaa rinnakkain. Kehittyneimmät suorittimet pystyvät käsittelemään tuhansia tehtäviä samanaikaisesti. (Järvinen, 2023, ss. 185–187)

5.2 Tekoälyn käytön riskit

Tekoäly ei ole ylivoimainen, vaikka valtamediasta helposti saa sellaisen kuvan. Sillä on heikkoutensa nopeasta kehityksestä huolimatta. (Järvinen, 2023, s. 177)

Ihmisen aivot ovat vertaansa vailla. Ne kehittyvät siinä, mihin niitä käytetään, mukautuvat iän tuomiin muutoksiin ja jopa paranevat sairauksista. Tekoälyllä ei ole mahdollisuuksia vastaavaan. Kun tekoälyn käyttämät tekniikat on valittu, ei se itse pysty niitä muuttamaan, kuvia tunnistava tekoäly ei opi robottiauton tekoälyksi. Heikko tekoäly voi siis toimia vain siinä tarkoituksessa, johon se on kehitetty. (Järvinen, 2023, ss. 177–178)

Tekoälyyn ja sen kehittämiseen kuluva energia ja sen ympäristövaikutukset puolestaan eivät ole saaneet paljoakaan palstatilaa mediassa. Suuret tekoälymallit kuluttavat paljon sähköä. Niissä on miljardeja parametreja, joita lasketaan moneen kertaan uudestaan, niin pitkään, kunnes saadaan haluttuja lopputuloksia. Kehittäjät haluavat tuottaa mahdollisimman

näyttävän tekoälymallin, eikä energiansäästöön ole kiinnitetty riittävästi huomiota, vaan malleissa on paljon turhaa ja päällekkäistä laskentaa. Vuonna 2019 tehdyssä Massachusettin yliopiston tutkimuksessa tutkittiin tavanomaisen syväoppimista hyödyntävän kielimallin aiheuttamia päästöjä. Järvisen mukaan tutkimuksessa arvioitiin, että pelkästään kielimallin kouluttaminen tuottaa yhtä paljon hiilidioksidipäästöjä kuin seitsemäntoista keskivertoamerikkalaista vuodessa. Kielimallin käytöstä aiheutuvat päästöt ovat koulutukseen verrattuna moninkertaiset. (Järvinen, 2023, ss. 184–185)

Ihmisellä on monia taitoja, joiden puuttuminen voidaan katsoa tekoälyn heikkoudeksi. Ensinnäkin ihmisellä on kyky tehdä päätöksiä yhdistämällä niin tietoista ajattelua kuin alitajuntaa. Tietoisesti analysoimme eri vaihtoehtoja, mutta alitajuisesti käsittelemme ympäristön mikrosignaaleja ja ne toimivat yhteistyössä hyvin. Työtehtävissä kertyy hiljaista tietoa, jota on mahdotonta sanoittaa tai kirjoittaa ohjeiksi uudelle työntekijälle eli tietoa ei voida siirtää myöskään tekoälylle. Ihminen pystyy toimimaan myös yllättävissä tilanteissa, toisin kuin tekoäly. Se käyttää ennalta annettuja kaavoja, joista valita, mutta ei kuitenkaan pysty soveltamaan tai muokkaamaan niitä yllättävän tilanteen tullen. Tekoäly ei tee virheitä, mutta se ei myöskään pysty ohittamaan muiden tekemiä virheitä, ainakaan ilman erillistä ohjelmointia, vaan se lopettaa toiminnan tai toimii virheellisesti, jos saa vääriä syötteitä. Viimeisenä, useimmiten asiakkaat haluavat asioida ihmisen kanssa. (Järvinen, 2023, ss. 188–194)

5.3 Tekoäly ja luonnollisen kielen käsittely käytännössä

”Kone on huono asioissa, jotka ovat helppoja ihmiselle, ja toisaalta hyvä asioissa, jotka ovat vaikeita ihmiselle.” Tämä muistisääntö kuvaa NLP:n käyttötapoja ja hyödyllisyyttä. Ihminen on hyvä luetun ymmärtämisessä, kaksoismerkityksissä, keskustelun käymisessä ja syy-seuraussuhteiden ymmärtämisessä. Kone puolestaan on ylivertainen suurten tekstimassojen käsittelyssä, sentimenttien tunnistamisessa tai poikkeamien havaitsemisessa suuresta aineistosta. Jotta kone on näissä hyvä, vaatii se sen opettamista datalla, jossa on oikea konteksti. Esimerkiksi kone, joka on opetettu lakitieteellisellä aineistolla, ei toimi parhaalla mahdollisella tavalla saadessaan tulkittavaksi puutarhanhoitoon liittyvää tekstiä. (Kananen & Puolitaival, 2019, s. 144)

Luonnollisen kielen luokittelua voidaan hyödyntää esimerkiksi tekstimuotoisten palautteiden käsittelyyn ja analysointiin. NLP:n avulla voidaan tuottaa tietoa palautteiden sävystä eli jakaa palautteet positiiviseen ja negatiiviseen palautteeseen tai tunnistaa palautteista eri merkityksiä eli semanttisuuksia. Palautteiden jaottelu aihealueittain tai kääntäminen eri kielille on myös mahdollista luonnollisen kielen käsittelyllä. Näillä erilaisilla käsittely- ja analysointitekniikoilla, saadaan lähes reaaliaikaisesti liiketoiminnan kannalta tärkeää tietoa niin kehityskohteista kuin hyvällä tasolla olevista aiheista. (Kananen & Puolitaival, 2019, s. 142)

NLP tarjoaa monia hyötyjä liiketoiminnan tarpeisiin. On kuitenkin tärkeää huomioida, että NLP ei ole virheetön. Käyttökohteen mukaan tulee harkita tuoko NLP:n käyttö enemmän hyötyä kuin haittaa ja minkälaisia NLP:n virheitä voidaan sietää. Esimerkkinä voidaan käyttää hakukoneita, jotka hakevat suuresta tietomäärästä nopeasti ihmisen tarvitseman tiedon, vaikka kaikki hakutulokset eivät välttämättä olekaan olennaisia. Toisaalta esimerkiksi lakitieteellisessä kontekstissa vastaavat virheet voivat olla sietämättömiä. Jotta NLP:tä voidaan ylipäänsä hyödyntää, se tarvitsee suuria määriä dataa. Usein konetta opetetaan ohjatun oppimisen keinoin, jolloin dataa tulisi ensinnäkin olla riittävästi ja toiseksi sen tulisi olla valmiiksi luokiteltua, mihin täytyy varata aikaa ja resursseja. (Kananen & Puolitaival, 2019, ss. 145–146)

5.4 Vastuullisuus tekoälyn käytössä

Tekoälyn liittyä useita eettisiä ongelmia, joista monet ovat samoja automaation ja robotiikan etiikan kanssa. Teknologioiden kytkeytyminen toisiinsa tuo esiin uusia näkökulmia eettisiin ongelmiin ja tekee niistä entistäkin tärkeämpiä huomioida. (Coeckelbergh, 2020/2021, s. 101) Tekoälyn vastuullisuutta ja eettisyyttä pohdittaessa on hyvä pitää mielessä mihin sitä käytetään, millaiseen dataan sen tekemät päätökset perustuvat sekä millaisia seurauksia sen käytöllä on (Kananen & Puolitaival, 2019, s. 222).

5.4.1 Yksityisyys

Usein tekoälyn käyttöön liittyen kerätään ja käytetään ihmisten henkilökohtaisia tietoja. Tekoälyn viitekehityksessä yksityisyyden ja aineiston suojaukseen liittyvät asiat ovat entistä

tärkeämpiä. Erityistä huomiota tulee kiinnittää aineiston käsittelyyn, joka tulee toteuttaa kunnioittaen yksilöiden yksityisyyttä sekä oikeutta tietää mitä heidän tiedoilleen tapahtuu. Toisaalta yksilöillä on myös oikeus kieltää omien tietojen käyttö sekä saada tietoonsa heistä kerätty aineisto. Kun tehdään esimerkiksi kyselyä, vastaajille voidaan selvittää aineiston käyttö ja pyytää lupa sen käsittelyyn, mutta aina tilanne ei ole näin selkeä. Esimerkiksi sosiaalisessa mediassa käyttäjät eivät välttämättä ole tietoisia mihin heistä kerättyä aineistoa käytetään. (Coeckelbergh, 2020/2021, ss. 101–102)

Henkilötietojen käsittelyssä on huomioitava yleinen tietosuoja-asetus, jolla suojataan yksityishenkilöiden oikeuksia henkilötietoihinsa. Asetuksella on myös yhtenäistetty henkilötietojen käsittelyä EU-alueella. (Euroopan parlamentti ja neuvosto, 2022) Mikäli henkilötietoja siirretään EU:n tai ETA-alueen ulkopuolelle, tietosuoja ei välttämättä ole vastaavalla tasolla mitä yleisellä tietosuoja-asetuksella on säädetty. Tämän vuoksi tietojen siirrolle EU:n tai ETA-alueen ulkopuolelle on määritetty tarkat edellytykset, joiden tulee täytyä, jos henkilötietoja aiotaan siirtää. (Tietosuojavaltuutetun toimisto, n.d.-a)

5.4.2 Vastuu

Jokainen ihminen on vastuussa niin tekemisistään kuin päätöksistään. Tekoälyn yleistyessä ja saadessa hoidettavakseen enemmän tehtäviä, joita ihmiset ovat hoitaneet, tulee pohdittavaksi, kenelle kuuluu vastuu. Kuka ottaa vastuun, kun jokin ei onnistu, tulee suuria vahinkoja tai tekoäly tuottaa hyötyä? Koska koneet eivät ole tietoisia tekemistään päätöksistä, täytyy vastuun olla ihmisellä olettaen, että hän on tietoinen siitä, mitä tekoäly tekee ja miksi tehtävään käytetään tekoälyä. Vastuu kysymykseen liittyy useita ongelmia, kuten tekoälyn nopea toiminta – kuinka ihminen voi ottaa vastuun päätöksestä, johon hänellä ei todellisuudessa ole mahdollisuutta vaikuttaa? Tekoälyn kehittämiseen ja käyttöön osallistuu useita henkilöitä – miten ja missä suhteessa vastuu jaetaan henkilöiden kesken? (Coeckelbergh, 2020/2021, ss. 113–119)

Vastuullisuuteen liittyy vahvasti myös selitettävyyys. Kun jokin menee pieleen, vaaditaan paitsi vastuuta myös selitystä tapahtuneelle. Joissakin tekoälysovelluksissa on selkeästi nähtävissä tapa, jolla lopputulokseen päädytään. Esimerkiksi päätöspuuta käyttävä tekoäly on ohjelmoitu niin, että päätös määrittyy syötteen perusteella, jolloin ihminen voi selittää tai

tekoälyä voidaan pyytää selittämään, miten päätökseen on päädytty. Kaikissa tekoälysovelluksissa päätökset eivät ole näin läpinäkyviä. Koneoppimista ja etenkin syväoppimista hyödyntävien tekoälyjen päätöksenteko ei ole yhtä lailla läpinäkyvää ja selitettävissä. (Coeckelbergh, 2020/2021, ss. 119–120) Syväoppimisen tekee päätöksen, mutta ei tuota ihmiselle ymmärrettävällä tavalla vastausta siitä, miten se on vastaukseen päätenyt. Tätä kutsutaan mustan laatikon ongelmaksi. (Raatikainen, 2021, ss. 181–182)

5.4.3 Vinouma

Vinoumalla tarkoitetaan tilannetta, jossa tekoäly päättyy ratkaisuun, joka voi olla epäoikeudenmukainen tietyille yksilöille tai ryhmille. Se on niin eettinen kuin yhteiskunnallinenkin ongelma, joka liittyy tekoälyyn perustuvaan datatieteeseen eikä se koske muita automaatioteknologioita. Epäoikeudenmukaisuutta ja vinoumia esiintyy myös yleisesti yhteiskunnassa, mutta tekoälyn kohdalla huolenaiheena on niiden jatkuminen ja vaikutusten kertaantuminen. (Coeckelbergh, 2020/2021, s. 127) Vinouma voi syntyä tekoälyn suunnittelun, testauksen tai soveltamisen yhteydessä. Esimerkiksi jos tekoälyn opetukseen käytettävä tietoaaineisto perustuu valkoihoisiin miehiin ja tekoälyä on tarkoitus hyödyntää koko väestöön, on lopputuloksena vinouma. Toisin sanoen tekoälyn päätökset ovat vinoutuneita, mikäli sille annettu tietoaaineisto on painottunut väärin. (Coeckelbergh, 2020/2021, ss. 130–131) Toisaalta kaikki maailmasta kerätty data on ainakin jossakin suhteessa vinoutunutta. Jos sitä lähdetään oikaisemaan, se ei enää ole totuudenmukaista, koska siinä luodaan uusia vinoumia. Näin päädytään filosofiseen ongelmaan, johon algoritmit tai ohjelmointi eivät tarjoa ratkaisua. (Järvinen, 2023, s. 180)

5.4.4 Datan laatu

Tekoälyn käytössä ja luonnollisen kielen käsittelyssä on otettava huomioon myös käytettävän datan laatu. Ohjatussa oppimisessa opetusdata tulee olla valmiiksi luokiteltua, mikä on usein työlästä. Luokittelu voidaan tehdä itse tai siihen voidaan hyödyntää kolmatta osapuolta, mutta luokittelun tekemiseen on olemassa myös työkaluja. Luokittelulla on suuri merkitys lopputulokseen. Ihmisten mielipiteet siitä miten data tulee luokitella, voivat erota toisistaan tai toimialalla voi olla käytössä yleinen standardi, joka puolestaan helpottaa luokittelua. Datan tulkinnan tulee vastata tavoitetilaa, minkä toteuttaminen voi olla hyvin

haastavaa. Joskus luokittelu voi vaatia toimialan erityistuntemusta tai data voi olla salassa pidettävää, jolloin luokittelua ei voida ulkoistaa. (Kananen & Puolitaival, 2019, ss. 145–147)

6 Avovastausten analysoinnin selvitys

Valtion talous- ja henkilöstöhallinnon palvelukeskus tuottaa valtionhallinnon virastoille, laitoksille ja rahastoille talous- ja henkilöstöhallinnon asiantuntijapalveluita.

Henkilöstötutkimus on yksi asiakkaille tarjottava palvelu, johon sisältyy järjestelmä tutkimuksen toteuttamiseen sekä tutkimusten tulosten virasto- ja valtiotasoinen raportointi. Henkilöstötutkimusta on lähdetty kehittämään vuoden 2019 hallitusohjelman linjauksien myötä. Tarkoituksena on uudistaa niin valtiotasoinen henkilöstötutkimuksen sisältö kuin järjestelmä, jolla tutkimuksia toteutetaan, mikä puolestaan mahdollistaa myös toimintatapojen kehittämisen ja uudistamisen. Uusi järjestelmä henkilöstötutkimuksen toteuttamiseen ja raportointiin on otettu pilotoitavaksi syksyllä 2023. Opinnäytetyö keskittyy tutkimuksen avoimien tekstivastausten käsittelyyn ja raportointiin uudessa järjestelmässä.

6.1 Selvityksen tavoite

Avoimet tekstivastaukset täydentävät muilta osin määrällistä tutkimusta, jossa vastausvaihtoehdot ovat asteikolla 1–5. Numeerisista vastauksista voidaan laskea keskiarvoja ja jakaumia, muodostaa hajontakaavioita ja trendejä, mutta avoimet vastaukset parhaimmillaan vastaavat kysymykseen 'miksi?' ja antavat perusteluita laskennallisille tuloksille. Henkilöstötutkimuksen palvelun tuottajana Palkeille on muodostunut käsitys siitä, että avoimien tekstivastausten käsittely asiakasvirastoissa vie paljon aikaa ja käsittelyn sekä raportoinnin lopputulos riippuu paljon työn tekevästä henkilöstä. Opinnäytetyössä on tarkoitus kerätä asiakkailta kokemuksia nykyisestä avovastausten käsittelystä sekä tarpeita avovastausten käsittelyn kehittämiseksi, selvittää millaisia ratkaisuja avovastausten käsittelyyn tällä hetkellä on saatavilla ja mitä hyötyjä tai riskejä tekoälyn hyödyntäminen avovastausten käsittelyssä tuo Palkeiden tuottamaan palveluun. Näitä näkökulmia yhdistelemällä syntyy opinnäytetyön tulos, jonka avulla tekoälyratkaisua avovastausten käsittelyyn voidaan mahdollisesti lähteä edistämään huomioiden niin asiakasnäkökulma, hyödyt ja riskit. Opinnäytetyön tuloksia voidaan käsitellä Palkeissa sisäisesti ja hyödyntää tekoälyn käyttöönotossa.

Opinnäytetyön teoriaosuudessa käsitellään niin henkilöstötutkimuksen kuin tekoälyn perusteita. Tietoperustassa selvitetään myös tekoälyä käytännössä luonnollisen kielen

käsittelyn alueella, mikä tuo pohjatietoa sille, miten tekstimuotoisia vastauksia on mahdollista käsitellä tekoälyn avulla. Henkilöstötutkimuksen tietoperustassa keskitytään selvittämään mitä henkilöstötutkimuksella tavoitellaan ja miten se toteutetaan. Jotta tekoälyn hyötyjä ja riskejä voidaan ymmärtää, käsitellään teoriaosuudessa tekoälyä perusteista käytäntöön saakka; mitä tekoäly on ja millaisiin tekniikoihin se voi perustua sekä miten tekoälystä on mahdollista tuottaa hyötyä henkilöstötutkimuksen osaksi. Teoriaosuudessa on selvitetty myös tekoälyn käytössä huomioitavia seikkoja, niin mahdollisuuksia kuin riskejäkin.

Opinnäytetyön aineistoa kerättiin asiakasorganisaatioista sekä järjestelmien toimittajilta. Asiakasorganisaatioiden henkilöstöhallinnon edustajille eli henkilöstöhallinnon ja henkilöstötutkimusten asiantuntijoille kohdennettiin verkkolomakekysely, joka koostui määrällisistä ja laadullisista kysymyksistä. Kyselyn tavoitteena oli kerätä tietoa avovastausten käsittelyn nykytilasta mahdollisimman laajasti eri asiakasorganisaatioissa. Lisäksi yksittäisille kyselyyn vastanneille järjestettiin haastattelut, joissa kerättiin syvemmin laadullista tietoa asiakasorganisaatioiden näkemyksistä ja tarpeista avovastausten käsittelyssä. Haastattelut toteutettiin osittain strukturoituna haastatteluna.

Asiakasnäkökulman lisäksi selvitettiin ratkaisujen ja järjestelmän toimittajilta, mitä he voivat tarjota asiakaskyselyn tarpeisiin. Markkinoilla olevien ratkaisujen toimittajille tehtiin sähköpostikysely, jolla pyrittiin selvittämään, millaisia ratkaisuja suomenkielisille avointen tekstivastausten käsittelylle on tällä hetkellä saatavilla ja millaisia tuloksia ratkaisut tuottavat. Tietoja saatavilla olevista ratkaisuista kerättiin myös toimittajien verkkosivuilta. Lisäksi nykyisen järjestelmän toimittajan kanssa järjestettiin avoin haastattelu, jossa toimittaja esitteli nykyisen järjestelmän tarjoamia tekoälyn mahdollisuuksia.

Opinnäytetyön tulosten perusteella on tarkoitus edistää valtionhallinnon digitalisaatiota käyttöönottamalla tekoäly avovastausten käsittelyyn, mikä osaltaan täydentää valtion yhteisen henkilöstötutkimuksen tuloksia. Tällä hetkellä valtion yhteisestä henkilöstötutkimuksesta raportoidaan keskiarvotietoja sekä organisaatiotasolla että valtiotasolla. Avovastauksia raportoidaan ainoastaan organisaatiotasolla, mutta valtiotasoisia yhteenvetotietoja avovastauksista ei ole saatavilla. Valtion yhteisen henkilöstötutkimuksen kohderyhmänä ovat kaikki valtion virastoissa, laitoksissa tai

rahastoissa työskentelevät henkilöt, joita on yhteensä noin 70 000. On siis perusteltua, että koontia näin suuren vastaajamäärän avoimista tekstivastauksista ei ole tehokasta toteuttaa ilman automaatiota tai tekoälyä. Valtiotasoiset avovastaukset tukisivat päätöksen tekoa hallituksen tasolla, kun valtion henkilöstöhallintoa halutaan kehittää paremmaksi.

6.2 Selvityksen toteutus ja tulokset

Tietoa avovastausten analysointiin liittyen tutkittiin kahdesta näkökulmasta – asiakkailta ja järjestelmän toimittajilta. Asiakasorganisaation edustajilta kerättiin tietoa tarpeista, joihin pyritään vastaamaan, yleisesti ratkaisujen toimittajilta kerättiin tietoa markkinoilla olevista mahdollisuuksista sekä nykyiseltä järjestelmän toimittajalta tietoa tarkemmalla ja teknisemmällä tasolla mahdollisuuksista tekoälyn hyödyntämiseen.

6.2.1 Asiakaskyselyn ja -haastattelujen suunnittelu ja toteutus

Palkeiden toiminnassa asiakkaiden tarpeet ovat keskiössä, joten myös henkilöstötutkimuksen avovastausten käsittelyn osalta haluttiin kuulla asiakkaiden näkemyksiä ja kokemuksia aiheesta. Asiakaskysely suunniteltiin niin, että toimeksiantaja saa asiakkailta, eli mahdollisen tulevan toiminnallisuuden käyttäjiltä, tarpeellista tietoa avovastausten käsittelyn nykytilanteesta asiakasorganisaatioissa. Jotta on mahdollista arvioida tekoälyllä tavoiteltavan hyödyn määrä, selvitettiin asiakkailta, kuinka paljon vastaajia organisaation henkilöstötutkimuksessa on, kuinka paljon työaika avointen tekstivastausten käsittelyyn kuluu ja millä tavoin avoimia tekstivastauksia organisaatioissa hyödynnetään. Kysely haluttiin pitää mahdollisimman tiiviinä, jotta vastaajia saataisiin useista organisaatioista. Tämän vuoksi kyselyyn otettiin organisaationperustietojen lisäksi yhdeksän kysymystä, joilla kerättiin tietoa nimenomaan avovastausten käsittelyn nykytilanteesta. Kyselyn lisäksi haastateltiin muutamaa kyselyyn vastannutta organisaation edustajaa. Haastatteluissa kerättiin tarkemmalla tasolla tietoa avovastausten käsittelystä, haasteista sekä tarpeista.

Kysely asiakasorganisaatioille toteutettiin verkkokyselylomakkeena, jonka linkki sekä saatekirje toimitettiin sähköpostiviestillä asiakasorganisaatioiden henkilöstöhallinnon edustajille toimeksiantajan viestintäosastolta. Saatekirjeessä kerrottiin, miksi kysely tehdään,

mihin vastauksia käytetään ja miksi vastaaminen on tärkeää. Kysely oli avoinna 11.1.2024–19.1.2024 eli 9 vuorokautta ja vastausaika sijoittui kahdelle eri viikolle, jolloin kyselyn saateviesti tavoitti mahdollisimman laajasti vastaajia. Vastauksien määrää tarkasteltiin vastausaikana muutaman kerran mahdollisen muistutusviestin lähettämistä varten. Vastaajia oli kuitenkin kertynyt kattavasti jo hyvissä ajoin ennen kyselyn päättymistä, joten muistutusviestiä ei lähetetty laajan vastaanottajajoukon vuoksi.

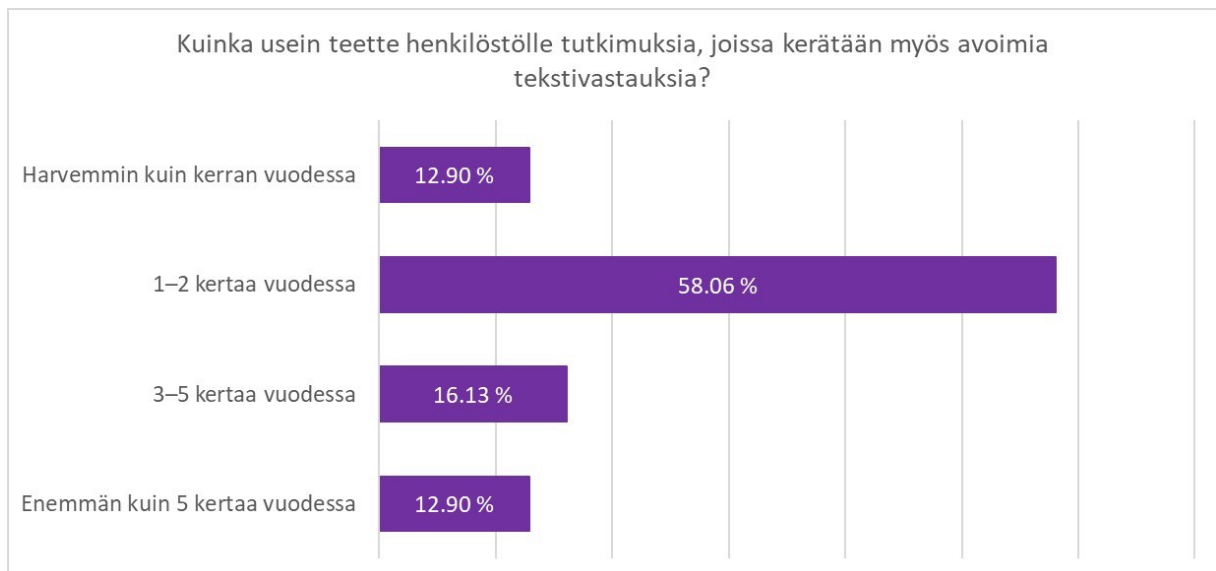
Vastausajan päätyttyä vastaukset vietiin Excel-taulukkoon, jossa jokainen vastaus oli omalla rivillään. Vastausten käsittely aloitettiin käymällä aineisto läpi silmäilemällä yleisesti vastauksia. Kysymyksistä kolmella oli selkeät vastausvaihtoehdot ja lisäksi yhden kysymyksen vastaukset olivat sellaisia, joista oli mahdollista muodostaa kaavio. Näistä neljästä määrällisestä kysymyksestä tehtiin Pivot-kaaviot Excelissä ja muotoiltiin ne sopivaksi. Kahden vaihtoehdon kaavioihin valittiin ympyräkaavio ja neljän vastausvaihtoehdon kaavioihin palkkikaavio, jotta luettavuus pysyisi mahdollisimman selkeänä. Kaavioissa käytettiin yhtenäisiä värisävyjä. Yksi kysymyksistä koski avovastauksien käsittelyyn käytettävää aikaa organisaatioissa. Tämän kysymyksen kohdalla vastaajien antamat arviot muutettiin tunneiksi, minkä perusteella laskettiin keskiarvoja, mediaaneja sekä vastausten pienimpiä ja suurimpia arvoja. Neljä muuta kysymystä olivat avovastauksia eli laadullisia kysymyksiä. Näiden käsittelyssä päästiin kokeilemaan, minkälaista avovastausten käsittely manuaalisesti on. Vastauksia käsiteltiin kysymyksittäin ja jokaisen vastauksen pääasiat purettiin omille riveilleen niin, että vastausten eri osille oli mahdollista lisätä asiasanoja. Asiasanat koottiin vastauksissa toistuvien teemojen mukaan ja vastauksiin lisättiin sopivat asiasanat jokaiseen vastauksesta erotettuun pääasiaan.

Kyselyn jälkeen haastateltiin neljää kyselyyn vastannutta asiakasorganisaation edustajaa. Haastateltavat valittiin vastausten perusteella niin, että mukaan saatiin henkilöstömäärältään eri kokoisia organisaatioita sekä erilaisia henkilöstötutkimuksen avovastausten käsittelytapoja. Haastattelut toteutettiin videohaastatteluina, osittain strukturoituna laadullisena haastatteluna, jonka runkoon valittiin seitsemän kysymystä. Haastattelussa otettiin myös huomioon haastateltavan esiin tuomat seikat, joihin pureuduttiin tarvittaessa tarkemmin lisäkysymyksillä.

6.2.2 Asiakaskyselyn tulokset

Asiakaskyselyyn vastasi yhteensä 31 henkilöä ja vastaajia oli 29:stä eri organisaatiosta. Valtion virastoja ja laitoksia on yhteensä 184, joten vastausaktiivisuudeksi virastotasolla saadaan noin 16 %. Kyselyyn vastanneiden organisaatioiden henkilöstön lukumäärä vaihteli muutamasta kymmenestä useaan tuhanteen. Keskiarvo organisaatioiden koolle oli noin 700 työntekijää. Henkilöstötutkimuksia tehdään vähintään kerran vuodessa suurimmassa osassa organisaatioista. Noin 13 prosenttia vastanneista tekee (Kuva 6) henkilöstötutkimuksia harvemmin kuin kerran vuodessa. Selkeästi yleisimmin, lähes 60 prosentin vastaajamäärällä, henkilöstötutkimuksia tehdään 1–2 kertaa vuodessa.

Kuva 6 Henkilöstötutkimusten määrä vuodessa



Kyselyssä pyydettiin arvioimaan työaika, joka koko organisaatiossa kuluu avovastausten käsittelyyn. Vastauksissa korostuu, että avovastausten käsittelyyn kuluva aika kokonaisuudessaan on vaikea arvioida, koska niitä käsitellään organisaatiossa monella tasolla ja eri henkilöiden toimesta. Osa vastaajista on arvioinut henkilöstöhallinnon käyttämää työaika ja osa koko organisaatiossa käytettyä työaika. Arviot, joissa on numeerisesti esitetty arvio käytetystä ajasta, muutettiin tuntimuotoiseksi laskennallisten tulosten muodostamista varten. Vastauksissa, joissa työmäärää ei arvioitu numeerisesti, ei ole huomioitu laskennallisissa tuloksissa. Näissä vastauksissa kuitenkin nousee esille, että avovastausten käsittelyyn kuluu huomattavan paljon työaika, joten olisi todennäköistä, että

näiden numeerinen arviointi olisi nostanut keskiarvoja. Arviot kuluvaista työajasta vaihtelivat yhdestä työtunnista yli tuhanteen työtuntiin, ja keskiarvoksi arvioidulle työmäärälle muodostui noin 80 työtuntia. Mediaaniarvo vastauksista oli noin 18 tuntia, eli yksittäiset vastaukset nostivat keskiarvoa huomattavasti. Kun laskelmiin huomioidaan organisaation henkilöstömäärä, voidaan laskea esimerkkejä avovastauksiin kuluvaista ajasta erikokoisissa organisaatioissa. Kuten taulukosta (Taulukko 1) nähdään, esimerkiksi 600 henkilön virastossa keskiarvo avovastausten käsittelyyn kuluvaista työajasta on koko organisaatiossa yhteensä 87 tuntia. Tässä on huomioitava avovastauksissa esiin noussut vastausten käsittely usealla organisaatiotasolla eli työaika ei ole yhden henkilön avovastausten käsittelyyn käyttämää työaikaa.

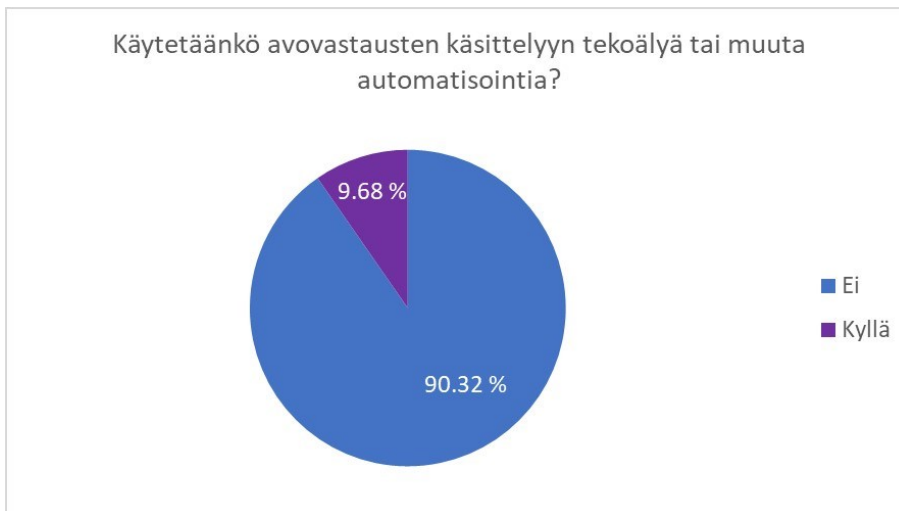
Taulukko 1 Organisaation kokoon suhteutettu työmäärän keskiarvo

Organisaation henkilömäärä	Keskiarvo työaika (h)
50	7,25
600	87
2900	420,29

Henkilöstötutkimuksista saatuja avovastauksia käsitellään kyselyyn vastanneissa organisaatioissa pääsääntöisesti manuaalisesti ja niitä käydään läpi usealla eri organisaation tasolla. Manuaalista työtä avovastausten käsittelyssä muodostuu henkilöstöhallinnon tai esihenkilöiden käydessä vastaukset läpi ja tehdessä niistä koosteista, joista yksittäisiä vastaajia ei voida tunnistaa sekä samalla siivotaan epäasialliset kommentit pois. Koosteita tehdään ryhmittelemällä ja luokittelemalla vastauksia usein toistuviin kategorioihin ja teemoihin. Vastauksissa korostuu, että vastauksia sellaisenaan ei toimiteta henkilöstön luettavaksi, vaan organisaation henkilöstöhallinnon avointen tekstivastausten läpikäynti on välttämätöntä ennen esittelyä koko henkilöstölle. Yleiselle tasolle tehdyt koosteet käydään useissa organisaatioissa henkilöstön kanssa läpi ja esiin nousseista asioista keskustellaan. Henkilöstöhallinnon asiantuntijoiden lisäksi organisaatioiden johto sekä yksiköiden johtajat saavat monissa organisaatioissa vastaukset luettavakseen sellaisenaan. Myös työsuojeluvaltuutetut ja luottamusmiehet saavat avovastaukset osassa organisaatioista luettavaksi.

Suurimmassa osassa vastanneista organisaatioista avointen tekstivastausten analysointiin ei käytetä tekoälyä tai muuta automatisointia (Kuva 7). Noin kymmenen prosenttia vastanneista kertoi, että heidän organisaatiossaan on käytetty tekoälyä tai muuta automatisointia avovastausten tiivistämiseen tai kehittämiskohteiden ja vahvuuksien esiin nostamiseen.

Kuva 7 Tekoälyn tai automatisoinnin käyttö avovastausten käsittelyssä



Vaikka tekoälyä tai automaatiota ei suurimmassa osassa vastaajien organisaatioissa käytetä, lähes kaikki vastanneet kuitenkin tekevät avovastauksista koosteita tai raportteja (Kuva 8).

Kuva 8 Avovastausten koosteet



Pääsääntöisesti avovastauksista tehdään yhteenvetoja yleisellä tasolla eli koosteita vastauksissa toistuvista teemoista. Myös sanapilviä, kehitettävien ja hyvien asioiden koosteita sekä suoria lainauksia käytetään, mutta näiden yleisyys vastauksien perusteella jäi melko pieneksi. Vastauksissa nousi esiin, että koosteita tehdään esimerkiksi Excelissä tai PowerPointissa. Pieni osuus vastaajista kertoo ostavansa koosteet avovastauksista palveluna, mutta suurin osa tekee koosteet itsenäisesti omassa organisaatiossaan (Kuva 9).

Kuva 9 Koosteiden tekeminen itse vai ostaminen palveluna



Kokonaisuudessaan avoimia tekstivastauksia hyödynnetään laajasti organisaatioissa. Avovastauksista saatuja palautteita käytetään yhdessä numeeristen tulosten kanssa muun muassa organisaation toiminnan ja työhyvinvoinnin kehittämiseen sekä kehittämiskohteiden tunnistamiseen. Organisaatioiden johto ja työsuojelutoimikunta saavat palautteet tueksi omaan työhönsä, henkilöstöhallinto saa tietoa kehitettävistä kohteista ja lisäksi palautteet käydään yhdessä läpi koko henkilöstön kanssa.

6.2.3 Asiakashaastattelujen tulokset

Avovastauksia käsitellään asiakasorganisaatioissa haastattelujen perusteella pääsääntöisesti itse organisaation sisällä useilla eri organisaation tasoilla. Avointen tekstivastausten

käsittelyn tapa niin eri organisaatioissa kuin organisaation sisäisesti vaihtelee paljon. Osassa organisaatioista avovastauksia luokitellaan ja asiasanoitetaan järjestelmällisesti, ja niistä koostetaan esitysmateriaaleja eri kohderyhmille. Osa haastatelluista kertoo organisaation tilanteen tällä hetkellä tai aiemmin avovastausten analysoinnin ulkoiselta palveluntarjoajalta, joka hyödyntää analysoinnissa tekoälyä. Suurimmalla osalla haastatelluista ei kuitenkaan ole käytössä automatisointi tai tekoälyä avovastausten käsittelyyn. Haastateltujen organisaatioiden välillä on vaihtelua siinä ketkä organisaatiossa saavat avoimet tekstivastaukset sellaisenaan luettavaksi. Osassa organisaatioista avoimet tekstivastaukset jaetaan sellaisenaan ryhmien esihenkilöille saakka, kun osassa avoimia vastauksia sellaisenaan lukevat vain henkilöstöhallinto sekä ylimmän johdon jäsenet. Selvää kuitenkin on, että avoimia tekstivastauksia kaikissa haastatelluissa organisaatioissa avoimia tekstivastauksia lukevat useat henkilöt eri työroolien näkökulmasta.

Avoimia tekstivastauksia kertyy organisaatioissa luettavaksi paljon, useista kymmenistä sivuista useisiin satoihin sivuihin. Luettavien vastausten määrä on suuri. Haastatteluista nousee esille, että vastausten määrä vaihtelee suuresti organisaation sisällä. Toisissa yksiköissä ei välttämättä anneta juuri lainkaan avointa palautetta, kun toisissa yksiköissä palautteet voivat olla hyvin pitkiä. Vastaajan palautteen sävy nähdään myös vaikuttavana tekijänä avovastauksien määrään – tyytymättömämmät vastaajat kirjoittavat enemmän avointa palautetta.

Manuaalisessa itse tehtävässä avointen tekstivastausten käsittelyssä nähdään hyvänä organisaation tuntemus, jolloin pystytään helpommin tunnistamaan avovastauksissa esiintyviä ilmiöitä. Myös avovastausten jakaminen sellaisenaan esihenkilöille koetaan hyvänä, koska tällöin esihenkilöiden on mahdollista hyödyntää vastauksia ryhmän toiminnan kehittämiseen. Suurimpana haasteena avointen tekstivastausten käsittelyssä koetaan suuri työmäärä ja käsittelyyn kuluva aika, lisäksi yhteyksien tunnistaminen muihin henkilöstötietoihin koetaan työlääksi. Tyytymättömien henkilöiden antamat vastaukset koetaan myös haasteena, koska sisällössä voi olla epäasiallisuuksia, joiden lukeminen on raskasta. Lisäksi haastatteluissa nousi esiin, että vastaajan avoin vastaus voi herkästi saada liikaa painoarvoa, koska avovastausten analysoija tulkitsee vastauksia omista näkökulmistaan.

Haastateltavat kokevat, että tekoäly voisi tuoda parannusta avointen tekstivastausten käsittelyyn etenkin työajan säästön näkökulmasta. Toisena mainituimpana hyötynä haastatteluissa nousee esiin tekoälyn objektiivisuus, jolloin ihmisen oma tulkinta ja painotus eivät vaikuta avovastausten koosteeseen. Ajatuksia tekoälyn hyödyistä nousi myös tarvittavien toimenpiteiden ja eniten nousseiden teemojen tunnistamisesta, raporteista sekä tehtävien toimenpiteiden vaikuttavuuden arvioinnista. Tekoälyn käytölle tunnistaan haastattelujen perusteella myös riskejä, joista suurimpina esiin nousevat tietosuoja sekä luotettavuus. Avovastauksissa tunnistetaan olevan arkaluontoista tietoa, minkä vuoksi tietosuoja ja tietoturva nähdään riskinä, johon tulee kiinnittää erityistä huomiota. Luotettavuus puolestaan nousee esiin sekä haastateltavien omien kokemusten sekä uutisoinnin pohjalta. Avovastausten analyysin tuloksien tulee olla riittävän luotettavia, jotta niihin voidaan pohjata organisaation kehittämistoimia.

Haastattelujen perusteella kehittyneemmille raporteille avovastausten hyödyntämisen tueksi olisi organisaatioissa käyttöä. Avovastausten sisältöjen luokitteluun kategorioihin tai teemoihin nousee haastatteluissa toivotuimmaksi raportointimuodoksi, mutta myös sentimenttien tunnistaminen sekä tiivistelmä avovastauksista saa kannatusta. Sanapilvi puolestaan jakaa mielipiteitä, osa haastateltavista kokee sen hyödyllisenä ja osa toisaalta ei koe sen tuovan vastaavasti lisäarvoa kuin esimerkiksi kategorisointi. Haastatteluissa nousee esiin myös toiveita edistyneemmästä raportoinnista, jossa avovastauksien tietoja voitaisiin vertailla eri taustakysymyksiin mukaan, annettujen numeeristen vastausten mukaan tai toisista henkilöstöjärjestelmistä kerättyihin tietoihin.

Suurimmassa osassa haastattelussa mukana olleista organisaatioista ei ole järjestelmällisesti käytössä tekoälyä muissa prosesseissa, mutta kokeiluja ja suunnitelmia on kaikkien pöydällä ja tekoälystä vastaavia ryhmiä on useammassa organisaatiossa. Organisaatioissa on käytetty tekoälyä esimerkiksi chatboteissa, kääntämisen ja kirjoittamisen apuna sekä ydintoiminnan prosesseissa. Osa haastatteluun osallistuneista kertoi, että organisaatiossa on kokeiltavana generatiivinen tekoäly, joka tukee esimerkiksi kirjoitustyössä ja tiedonhankinnassa. Henkilöstöhallinnossa haastateltavat eivät tunnista käytössä olevia tekoälykohteita. Haastateltavat arvioivat henkilöstön mieltävän tekoälyn käyttöön liittyvää tietoturvaa ja tietosuojaa sekä suhtautuvan kriittisesti sen lähteisiin. Henkilöstön arvioidaan olevan kuitenkin avoimia ja innostuneita tekoälyn käyttöön, henkilöstö nostaa esiin myös kohteita,

joissa tekoälyä voitaisiin hyödyntää. Kaiken kaikkiaan haastattelujen perusteella organisaatioiden henkilöstö suhtautuu tekoälyn käyttöön innostuneesti varustettuna terveellä kriittisyydellä.

6.2.4 Toimittajakyselyn suunnittelu, toteutus ja päätelmät

Opinnäytetyön yhteydessä haluttiin selvittää, millaisia ratkaisuja tällä hetkellä on saatavilla etenkin suomenkielisten avointen tekstivastausten analysointiin ja kuinka niillä pystytään vastaamaan tunnistettuihin ja asiakkaiden esiin tuomiin tarpeisiin. Ratkaisujen toimittajille toteutettiin lyhyt sähköpostikysely, jossa selvitettiin, onko kyseisellä yrityksellä tarjolla ratkaisua, jolla voidaan käsitellä ja analysoida avoimia tekstivastauksia ja millaisia raportteja ratkaisulla voidaan tuottaa asiakkaille. Kyselyllä selvitettiin myös hyödyntääkö ratkaisu tekoälyä ja onko kyseessä on-premise-palvelu vai pilvipalvelu. Yrityksiä pyydettiin myös kertomaan ovatko he tunnistaneet haasteita ratkaisun käyttöönotossa ja miten näitä haasteita voitaisiin välttää.

Toimittajakysely lähetettiin kymmenelle toimittajalle, jotka valikoituivat mukaan verkkosivujen perusteella. Mukaan valittujen toimittajien verkkosivuilta löytyi tietoa, että heillä on mahdollisesti tarjolla kuvattu ratkaisu tai kokemusta henkilöstötutkimuksista. Vastauksia kyselyyn saatiin kolmelta toimittajalta eli otanta toimittajien osalta on melko kapea. Vastauksista koostettiin taulukko, johon koottiin tiedot vastauksista kysymyksiin. Koostetaulukoon haettiin tietoa myös toimittajien verkkosivuilta mahdollisilta osin. Suurella osalla toimittajista verkkosivuilla oleva tieto ei ollut niin tarkalla tasolla, jotta tietoja olisi voitu ottaa mukaan koosteeseen. Taulukkoon koottujen tietojen perusteella pystyttiin tekemään yhteenveto ratkaisuksista, joita on valmiina tuotteina tarjolla.

Pienestä vastaajamäärästä huolimatta selvityksestä kävi ilmi, että suomenkielisille avointen tekstivastausten analysoinnille on tarjolla valmiita ratkaisuja ja sovelluksia, jotka tuottavat erilaisia raportteja niin johdon kuin henkilöstöhallinnon käyttöön. Toimittajien tarjoamat raporttimallit avovastauksien koostamiseen lähtivät liikkeelle sanapilvi-tyyppisestä ratkaisusta, mutta myös sentimenttejä, asiasanoitusta ja kategorisointia hyödynnetään. Edistysellisempänä vaihtoehtona kahdella vastanneesta toimittajasta on käytössä

trendianalyysit, joissa tarkastellaan ajan kuluessa tapahtuneita muutoksia sekä trendien vaikutusta esimerkiksi työtyytyväisyyteen.

6.2.5 Nykyisen järjestelmän mahdollisuudet

Nykyisessä järjestelmässä ei vielä ole käytössä tekoälyä, mutta kyvykkyyksiä siihen on olemassa. Järjestimme nykyisen järjestelmän toimittajien kanssa vapaamuotoisen keskustelutilaisuuden, jossa kävimme ensimmäistä kertaa yhdessä läpi niin järjestelmän mahdollisuuksia kuin Palkeiden ajatuksia tekoälyn käyttöönotosta. Keskustelussa nousi esiin muutamia vaihtoehtoja, jotka voisivat olla mahdollisia askeleita tekoälyn hyödyntämiseen. Näitä olivat klusterointi, NLU sekä generatiivinen tekoäly.

Klusteroinnissa yhdistetään avovastauksissa nousseita asioita kategorioiksi. Klusterointi ei vaadi erillistä koulutusta käyttöönotettavaksi vaan toimii ohjaamattoman oppimisen kaltaisesti. Koska klusteroinnin mallia ei erikseen kouluteta, se ei osaa huomioida esimerkiksi eri kielillä annettuja vastauksia, vaan vastaukset tulisi kääntää samalle kielelle ennen klusterointia. Kuten ohjaamattomassa oppimisessä, voi osa tämän menetelmän kategorioista olla epätarkkoja tai aiheeseen liittymättömiä.

NLU-malli eli luonnollisen kielen ymmärtäminen on ohjatun oppimisen kaltainen malli. NLU-mallilla voidaan hakea avovastausaineistosta teemoja tai asiasanoja, jotka on ennalta määritetty koulutusaineistossa. Koska tämä malli on koulutettava ja koulutusdataa tarvitaan kymmeniätuhansia rivejä, on käyttöönottoon varattava huomattavasti työaika. Käytettävät kategoriat ja aihealueet on kerättävä manuaalisesti olemassa olevasta aineistosta ja merkittävät ne koulutusdataan. Toisaalta myös tulokset ovat tarkempia ja paikkaansa pitävämpiä, kun malli on koulutettu oikealla aineistolla, eli sille on annettu valmiiksi kategoriat tai aihesanat, joihin vastaukset luokitellaan. Koulutuksen jälkeen malli tuottaa tulokset oikeasta aineistosta nopeasti eikä tällöin työaika mene kategorioiden koostamiseen.

Keskustelussa nousi esiin generatiivisen tekoälyn käyttäminen avovastausten analysoinnissa tai kategorioiden muodostamisesta koulutusaineistoon. Generatiivisen tekoälyn etuna voidaan nähdä esimerkiksi nopea käyttöönotto, mutta heikkoutena tällä hetkellä on

epätarkkuus vastauksissa, kun kyseessä on suomenkielinen työtyytyväisyyteen painottuva aineisto. Muita keskustelussa esiin nousseita aiheita oli sentimenttien tunnistaminen avovastauksista, numeeristen vastausten tulkinta ja trendien ennustaminen sekä käyttäjää järjestelmän käytössä avustava tekoäly. Koska tapaaminen oli ensimmäinen laatuaan, aiheesta ei edetty tekniselle tai käytännön tasolle, mutta aiheesta otettiin ensimmäinen askel kehittämistä. Keskustelun pohjalta Palkeissa voidaan lähteä suunnittelemaan käyttötapauksia ja suuntaa, johon tekoälyn käyttöönotossa halutaan lähteä.

6.3 Tekoälyn mahdollisuudet avovastausten käsittelyssä

Niin teoriaosuuden tietojen, kyselyiden sekä haastattelujen perusteella voidaan todeta, että tekoäly tarjoaa uusia mahdollisuuksia avointen tekstivastausten käsittelyyn. Tekoäly tuo prosessiin paljon hyötyjä, mutta siihen liittyy myös useita huomioitavia asioita.

Keskeisimmät hyödyt tekoälyn käytössä henkilöstötutkimuksen avovastausten käsittelyssä liittyvät työajan säästöön sekä objektiivisuuteen. Tällä hetkellä organisaatioissa käytetään avovastausten lukemiseen ja kategorisointiin kymmeniä tai jopa satoja tunteja aikaa useilla organisaation eri tasoilla. Avoimia tekstivastauksia yhdestä valtion yhteisestä henkilöstötutkimuksesta kertyy henkilöstön määrän mukaan useita kymmeniä tai satoja sivuja. Näiden lukeminen, tulkitseminen ja eniten mainintoja saaneiden teemojen hahmottaminen on työlästä. Työajan säästön näkökulmasta tekoäly tuo selkeitä hyötyjä avovastausten käsittelyyn. Kun ihminen lukee tunteja avovastauksia ja merkitsee niihin kategorioita, niin tekoäly käsittelee vastaavan aineiston hetkessä.

Avovastauksia lukee samassa organisaatiossa useita henkilöitä, joista jokainen tekee tulkinnan omista lähtökohdistaan. Tämä voi vaikuttaa avovastausten tulkinnan lopputulokseen, koska eri henkilöille eri asiat voivat nousta aineistosta merkittäviksi. Ihmisen tulkintaan vaikuttaa myös avovastausten määrä ja niiden lukemiseen kuluva aika. Kun vastauksia luetaan pitkiä aikoja, ihmisen keskittyminen heikkenee, mikä puolestaan voi vaikuttaa avovastausten tulkintaan. Kaikilla ihmisillä on oma taustansa, joka väistämättä vaikuttaa tekstin tulkintaan, mutta tekoäly on objektiivinen. Tekoälyllä pystytään siis vastaamaan myös tähän haasteeseen, huomioiden käytettävä tekoälytekniikka. Esimerkiksi henkilöstötutkimuksen avovastauksiin koulutettu tekoäly palauttaa jokaisella kerralla samat

tiedot eikä tulkitse aineiston avovastauksia eri tavalla kellonajan mukaan. Tekoäly ei myöskään väsy eikä aineiston rivien määrä aiheuta keskittymisen herpaantumista, vaan pystyy tekemään olennaiset havainnot aineistosta.

Tekoälyn avulla on mahdollista tehdä henkilöstötutkimuksen avovastauksista vertailukelpoisia raportteja, joita voidaan hyödyntää myös pidemmällä aikavälillä esimerkiksi trenditietojen raportoinnissa. Tekoälyllä voidaan jakaa avovastausaineisto esimerkiksi ennalta määritettyihin kategorioihin ja raportoida näiden kategorioiden esiintymistiheyttä aineistossa sekä luokitella avovastauksia sentimentteihin. Jo näillä raporteilla organisaatioissa pystyttäisiin keskittymään olennaisiin teemoihin, joita avovastauksista on noussut. Avovastauksista saatavaa tietoa pystytään tekoälyn tuottamalla analyysillä ja raporteilla hyödyntämään nopeammin sekä ketterämmin. Tekoäly jatkaa koko ajan kehittymistään ja se tulee tarjoamaan tulevaisuudessa lisää mahdollisuuksia ja innovaatioita niin tekstinkäsittelyyn kuin raportointiin.

Haastatteluissa nousi esiin huomio, että tyytymättömät vastaajat antavat enemmän avovastauksia kuin tyytyväiset vastaajat. Tämä ilmiö lienee tuttu jokaiselle, huonosta palvelusta kerrotaan kymmenelle tuttavalle, kun hyvästä palvelusta mainitaan ehkä vain yhdelle. Se voidaanko tätä estää tekoälyn hyödyntämisellä, on mielenkiintoinen kysymys. Toisaalta voidaan ajatella, että tyytymättömän vastaajan palaute muuttaa kokonaisuutta negatiivisemmaksi, mutta toisaalta se kertoo avovastauksen sävystä. Yksi ratkaisu tähän olisi vertailla numeerisiin kysymyksiin tyytymättömästi vastanneiden avovastauksia numeerisiin kysymyksiin tyytyväisesti vastanneiden avovastauksiin.

Tietosuoja ja tietoturva nousevat suurimmaksi huomioitavaksi asiaksi avovastauksien käsittelyssä tekoälyn avulla. Henkilöstötutkimuksen avovastaukset sisältävät arkaluonteisia tietoja ja jokainen vastaaja voi kirjoittaa niihin mitä haluaa eikä tällöin avovastausaineiston sisältämistä tiedoista voida olla täysin varmoja ennen kuin ne on luettu läpi. Vaikka vastaajia voidaan ohjeistuksella kehottaa kirjaamasta tunnistetietoja avovastauksiin, voivat ne sisältää esimerkiksi henkilötietoja. Henkilötietoihin sovelletaan EU:n tietosuoja-asetusta, joka edellyttää, että tietoja käsitellään EU- ja ETA-alueen sisällä lainmukaisesti. Avovastaukset voivat sisältää myös arkaluonteisia tietoja henkilöstöstä tai organisaatiosta, mikä myös

osaltaan vaatii luottamuksellista ja asianmukaista tietojen käsittelyä myös tekoälyratkaisussa.

Toinen tärkeä huomioitava seikka on luotettavuus. Analysoituihin tietoihin tulee voida luottaa, jotta organisaatiot voivat käyttää tietoja toimintansa kehittämiseen.

Avovastauksissa on tärkeää tietoa, joka halutaan hyödyntää organisaatioiden toiminnan kehittämisessä. Avoimista tekstivastauksista täytyy poimia oikeat tiedot, jotta pystytään kehittämään oikeita asioita eikä resursseja käytetä väärin asioiden kehittämiseen. Myös henkilöstön on tärkeää voida luottaa avovastauksista saatuihin tuloksiin, jotta heidän on helpompi sitoutua organisaation kehitystoimenpiteisiin. Lisäksi on huomioitava, että valtion yhteistä henkilöstötutkimusta toteutetaan erilaisille organisaatioille, joilla on erilaisia tarpeita. Erilaisten tarpeiden huomioiminen on tärkeää, jotta asiakasorganisaatiot saavat käyttöönsä heidän tarpeisiinsa sopivia analyysejä avovastauksista.

Kolmantena huomioitavana asiana esiin nousee koulutusdatan merkitys. Jotta tekoälyratkaisu on mahdollista ottaa käyttöön, tulee koulutusdataa olla paljon. Lisäksi koulutusdataan tulee olla merkitty halutut kategoriat, joiden valitsemiseen tulee kiinnittää erityistä huomiota. Organisaatioissa avovastauksista voi nousta esille hyvin erityyppisiä asioita, eli kategorioihin on tunnistettava organisaatioille tärkeimmät asiat. Kategorioiden valinnassa ja koulutusdatan koodaamisessa on kiinnitettävä erityistä huomiota siihen, että aineisto on totuudenmukainen eikä koulutusvaiheessa pääse muodostumaan vinoumaa käytettävään tekoälyratkaisuun. Kategorioiden valitsemiseen ja koulutusdatan koodittamiseen on varattava riittävästi työaika, jotta on mahdollista saada käyttöön luotettava ja laadukas tekoälyratkaisu avovastausten analysointiin.

Tekoälyn käyttöönotto henkilöstötutkimuksen avovastausten analysointiin vähentäisi niiden käsittelyyn kuluva työaika. Työaika olisi tällöin mahdollista käyttää esille nousseisiin teemoihin, niiden juurisyihin ja toimenpiteisiin. Tekoälyn tekemät luokittelut ovat objektiivisempia kuin ihmisen tekemät luokittelut, kun panostetaan koulutusdatan laadukkaaseen valmisteluun. Tekoäly tulee olemaan osa organisaatioiden tulevaisuutta monella eri osa-alueella. Kuten kaikessa kehittämisessä, myös henkilöstötutkimuksien avovastausten raportoinnin ja analysoinnin osalta on hyvä pitää mielessä, että kehittämistä tehdään tiedolla johtamisen ja keskustelun tueksi, ei toisin päin.

7 Johtopäätökset ja pohdinta

Tekoäly on viime vuosina kehittynyt kovaa vauhtia eikä siinä näy hiipumisen merkkejä. Kehitteillä on jatkuvasti uusia menetelmiä ja käyttökohteita tekoälylle. Tekoäly on mukana jo lähes kaikkien ihmisten arkipäivässä, ja organisaatiot ottavat tekoälysovelluksia käyttöön eri prosesseihin. Myös valtionhallinnossa on tavoitteena edistää digitalisaatiota ja sitä kautta käyttöönottaa uusia työelämää helpottavia ratkaisuja. Luonnollisen kielen käsittely on yksi osa-alue, jolla kehitystä on tapahtunut paljon niin koneoppimisen puolella kuin generatiivisen tekoälyn osa-alueella. Englanninkielisiä ratkaisuja on ollut markkinoilla jo pidemmän aikaa, mutta myös suomenkielisiä kielimalleja on jo kehitetty ja kehitetään edelleen, mikä osaltaan mahdollistaa uusia käyttökohteita tekoälylle.

Valtion yhteisen henkilöstötutkimuksen on tarkoitus tuottaa tietoa työntekijäymmärryksestä sekä työhyvinvoinnista organisaatioiden johdolle, henkilöstöhallinnolle ja koko henkilöstölle. Henkilöstötutkimuksista kertyy suuria määriä avovastauksia, joiden tehokkaaseen hyödyntämiseen tarvitaan yhteenvetoja ja koosteita. Avovastausten läpikäymiseen, luokitteluun ja tulkintaan kuuluu organisaatioissa huomattavia määriä työaikaa, ja ihmisen tuottama kooste sisältää aina lukijan omaa tulkintaa. Tekoälyn avulla on mahdollista koostaa avovastauksista erilaisia raportteja ja analyysejä, kuten esimerkiksi kategorisointia ja sentimenttianalyysejä.

Opinnäytetyössä tehtiin verkkolomakekysely asiakasorganisaatioiden edustajille, sähköpostikysely markkinoilla oleville yrityksille, osittain strukturoitu haastattelu asiakaskyselyyn vastanneille asiakasorganisaatioiden edustajille sekä avoin haastattelu nykyisen järjestelmän toimittajille. Nämä menetelmät sopivat pääosin hyvin opinnäytetyöprojektiin ja niillä saatiin kerättyä tietoa niin asiakkaiden tarpeista kuin teknisistä mahdollisuuksista avovastausten käsittelyyn.

Asiakkaiden tämänhetkisestä avovastausten käsittelystä ja siihen kuluvaista työajasta saatiin paljon tietoa verkkokyselyllä. Myös asiakkaiden haastattelut toivat paljon näkemyksiä tekoälyn käytöstä ja avovastausten käsittelystä. Haastatteluissa nousi hyvin esiin organisaatioissa tunnistetut tarpeet avovastausten käsittelyyn sekä huomioitavat asiat tekoälyn käytössä. Markkinoilla olevilta ratkaisujen toimittajilta vastauksia kyselyyn saatiin

melko vähän, minkä vuoksi sen osalta tutkimuksen otanta jäi pieneksi. Kuitenkin kyselyn perusteella voitiin todeta, että ratkaisuja avointen tekstivastausten käsittelyyn löytyy useammalta toimittajalta. Nykyisen järjestelmän toimittajan kanssa pidetyssä tapaamisessa avovastausten käsittelyn mahdollisuuksista päästiin teknisemmälle tasolle, lopputuloksena, että myös nykyisessä järjestelmässä on kyvykkyksiä avovastausten käsittelyyn tekoälyn avulla.

Teorian, asiakaskyselyn ja -haastattelujen sekä järjestelmän toimittajilta kerätyn tiedon perusteella koostettiin huomioitavat asiat tekoälyn käyttöön otossa henkilöstötutkimuksen avovastauksiin. Koostettuja asioita tullaan käyttämään hyödyksi toimeksiantajan kehittäessä henkilöstötutkimuksen avovastausten raportointia ja analysointia. Toimeksiantaja oli tyytyväinen opinnäytetyön tuloksiin ja niitä tullaan esittelemään toimeksiantajan henkilöstölle.

8 Yhteenveto

Valtion yhteistä henkilöstötutkimusta on lähdetty kehittämään vuoden 2019 hallitusohjelman tavoitteen perusteella. Henkilöstötutkimus koostuu numeerisista väittämistä ja avoimista tekstivastauksista. Tässä opinnäytetyössä selvitettiin miten avointen tekstivastausten käsittelyä ja analysointia voidaan helpottaa tekoälyn avulla.

Opinnäytetyössä selvitettiin myös, millaisia tarpeita asiakkailla avovastausten käsittelyyn on ja millainen niiden käsittelyprosessi tällä hetkellä organisaatioissa toimii. Aineiston perusteella koostettiin opinnäytetyössä esiin nousseita mahdollisuuksia ja riskejä, jotka on hyvä huomioida, kun tekoälyä otetaan käyttöön henkilöstötutkimuksen avovastausten analysointiin.

Opinnäytetyön tutkimuskysymyksiin saatiin kattavasti aineistoa eri näkökulmista ja niihin on mielestäni vastattu hyvin. Tutkimuskysymyksiä on käsitelty niin teoreettisesta näkökulmasta kuin asiantuntijoiden kokemusten kautta. Opinnäytetyön tuloksia tullaan hyödyntämään toimeksiantajan henkilöstötutkimus järjestelmän kehitystyössä ja niitä tullaan käsittelemään toimeksiantajalla sisäisesti.

Olen oppinut opinnäytetyöprosessin aikana paljon niin tekoälystä kuin henkilöstötutkimuksesta, jotka ovat aiheina hyvin erilaiset – tekninen ja humaani. Vaikka tekoäly on läsnä jo lähes jokaisen suomalaisen arkipäivässä, ei sen taustalla olevia tekniikoita tule ajatelleeksi joka päivä. Myös henkilöstötutkimus vaikuttaa paljon jokaisen työntekijän arkeen, mutta harvoin tulee pohdittua missä kaikkialla ja miten siitä saatavia tietoja voidaan hyödyntää. Opinnäytetyö on kasvattanut mielenkiintoani oppia lisää niin tekoälystä kuin henkilöstötutkimuksista.

Lähteet

Brooks, R. (2021). *What is reinforcement learning?* University of York.

<https://online.york.ac.uk/what-is-reinforcement-learning/>

Camposato, O. (2020). *Artificial Intelligence, Machine Learning, and Deep Learning*. Walter De Gruyter, Inc.

Camposato, O. (2021). *Natural Language Processing Fundamentals for Developers*. Mercury Learning and Information.

Chishti, S., Bartoletti, I., Leslie, A., & Millie, S. M. (2020). *The AI Book: The Artificial Intelligence Handbook for Investors, Entrepreneurs and FinTech Visionaries*. Wiley.

Coeckelbergh, M. (2020). *Tekoälyn etiikka* (K. Pietiläinen, Käänt.). Terra Cognita.

Euroopan parlamentti ja neuvosto. (2022). *Yleinen tietosuoja-asetus (GDPR) | EUR-Lex*.

<https://eur-lex.europa.eu/Fl/legal-content/summary/general-data-protection-regulation-gdpr.html>

Great Place To Work. (2023). Työhyvinvointi – kulmakivet hyvinvoivaan työyhteisöön. *Great Place To Work*. <https://greatplacetowork.fi/artikkelit/tyohyvinvointi/>

Günther, K., & Hasanen, K. (n.d.-a). *Tyypittely*. Tietoarkisto.

<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/tyypittely/>

Günther, K., Hasanen, K., & Juhila, K. (n.d.-a). *Johdanto: Analyysi ja tulkinta*. Tietoarkisto.

<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/analyysi-ja-tulkinta/>

Huhta, M., & Myllyntaus, V. (2021). *Työnantajabrändi ja työntekijäkokemus*. Alma Talent Oy.

Hyppänen, R. (2013). *Esimiesosaaminen. Liiketoiminnan menestystekijä* (3. p.). Edita Publishing Oy.

Itä-Suomen yliopisto. (2020). *Parempaa data-analyysiä tehokkaammilla klusterointimenetelmillä*. Itä-Suomen yliopisto. <https://www.uef.fi/fi/artikkeli/parempaa-data-analyysia-tehokkaammilla-klusterointi-menetelmilla>

Joki, M. (2021). *Henkilöstöasiantuntijan käsikirja*. Kauppakamari.

Juhila, K. (n.d.-a). *Koodaaminen*. Tietoarkisto.

<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/koodaaminen/>

- Juhila, K. (n.d.-b). *Teemoittelu*. Tietoarkisto.
<https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/teemoittelu/>
- Juuti, P., & Puusa, A. (2020). *Laadullisen tutkimuksen näkökulmat ja menetelmät*. Gaudeamus.
- Järvinen, P. (2023). *Tekoäly ja minä: Ihmisenä tekoälyn aikakaudella*. Tammi.
- Kananen, H., & Puolitaival, H. (2019). *Tekoäly: Bisneksen uudet työkalut*. Alma Talent Oy.
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3), 3713–3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Klemetti, A. (2022, helmikuuta 24). Tekoälyllä kohti ketterää digisalamointia. *Robologi*.
<https://blogit.metropolia.fi/robologi/2022/02/24/tekoalylla-kohti-ketteraa-digisalamointia/>
- Kolari, J., & Kallio, A. (2023). *Tekoäly 123*. Docendo.
- Merilehto, A. (2018). *Tekoäly: Matkaopas johtajalle*. Alma Talent Oy.
- Nelson, E. B. (2019). *AI Concepts for Business Applications*. Business Expert Press.
- Raatikainen, P. (2021). *Tekoäly, ihminen ja yhteiskunta: Filosofisia näkökulmia*. Gaudeamus.
- Ropponen, A. (2022). *Aihemallinnus ja muut menetelmät*.
https://wiki.eduuni.fi/download/attachments/266427730/Bibliometrikkaseminaari2022_Ropponen.pdf?version=1&modificationDate=1649765940283&api=v2
- Salonen, A. O., Lehtinen, E., Helajärvi, H., Sahimaa, J., Hautala, J., Viljamaa, J., Mikkola, J., Huitti, M., Huutilainen, M., Remes, S., Routarinne, S., & Dunderfelt, T. (2022). *Paremmän työelämän opas*. Minea.
- Saramies, J., & Törnroos, M. (2021). *Henkilöstöanalytiikka: Mittaa, ymmärrä, menesty*. Alma Talent Oy.
- Sosiaali- ja terveysministeriö. (n.d.). *Työhyvinvointi*. Sosiaali- ja terveysministeriö.
<https://stm.fi/tyohyvinvointi>
- Tietosuojavaltuutetun toimisto. (n.d.-a). *Henkilötietojen siirrot Euroopan talousalueen ulkopuolelle*. Tietosuojavaltuutetun toimisto. <https://tietosuoja.fi/henkilotietojen-siirrot-etan-ulkopuolelle>
- Työturvallisuuskeskus. (n.d.). *Työhyvinvointi*. Työturvallisuuskeskus.
<https://ttk.fi/tyoturvallisuus/tyohyvinvointi/>

- Valtion talous- ja henkilöstöhallinnon palvelukeskus. (n.d.-c). *5. Raportointi*. Haettu 27.7.2023 osoitteesta <https://www.vmbaro.fi/ohjeet/index.php/5-raportointi>
- Valtion talous- ja henkilöstöhallinnon palvelukeskus. (n.d.-f). *Henkilöstötutkimusjärjestelmän (VMBaro) toimittajaksi on valittu Sofigate Oy*.
<https://palkeet.fi/uutiset/henkilostotutkimusjarjestelman-vmbaro-toimittajaksi-on-valittu-sofigate-oy/>
- Valtion talous- ja henkilöstöhallinnon palvelukeskus. (n.d.-d). *Palkeet*.
<https://www.palkeet.fi/palkeet.html>
- Valtion talous- ja henkilöstöhallinnon palvelukeskus. (n.d.-e). *Palvelumme*. *Palkeet*.
<https://palkeet.fi/palvelut/>
- Valtion talous- ja henkilöstöhallinnon palvelukeskus. (n.d.-b). *Työtyytyväisyystutkimus*.
Haettu 25.7.2023 osoitteesta <https://www.vmbaro.fi/tyoetyytyvaeisyystutkimuksen-uudistus-2016>
- Valtion talous- ja henkilöstöhallinnon palvelukeskus. (n.d.-a). *Valtion henkilöstötutkimus ja Tutka*. *Palkeet*. <https://palkeet.fi/palvelut/valtion-henkilostotutkimus-ja-tutka/>
- Valtioneuvosto. (2019). *Pääministeri Sanna Marinin hallituksen ohjelma 10.12.2019*.
https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161931/VN_2019_31.pdf?sequence=1&isAllowed=y
- Valtiovarainministeriö, Happonen, M., & Laine, M. (2023). *Henkilöstöjohtamisen uudistusohjelma*. https://api.hankeikkuna.fi/asiakirjat/45e248d2-ca32-4a20-ade8-e809a2af7ecb/5d84140f-f80c-4161-ba5c-e1b61dd94fcb/KIRJE_20230329080958.PDF

Liite 1: Aineistonhallintasuunnitelma

Opinnäytetyön yhteydessä toteutetun asiakaskyselyn vastauksia säilytetään toimeksiantajan kyselyn toteutusjärjestelmässä sekä toimeksiantajan omistamalla verkkolevyllä kuusi kuukautta asiakaskyselyn päättymisestä lukien. Tämän jälkeen tiedot hävitetään tietoturvallisesti.

Asiakaskyselystä tehty kooste säilytetään toimeksiantajan omistamalla verkkolevyllä vähintään vuoden ajan. Tietoja käytetään tarvittaessa toimeksiantajan kehityskohteisiin myöhemmin.

Asiakashaastattelujen Teams-nauhoitteet ja litterointitiedostot säilytetään toimeksiantajan omistamalla verkkolevyllä kuusi kuukautta haastatteluista lukien. Haastatteluista tehty kooste säilytetään toimeksiantajan omistamalla tietokoneella vähintään vuoden ajan opinnäytetyön hyväksymisestä. Tietoja käytetään tarvittaessa toimeksiantajan kehityskohteisiin myöhemmin.

Järjestelmän toimittajien vastaukset sähköpostikyselyyn sekä kooste vastauksista säilytetään opiskelijan tietokoneella ja niistä tehdään varmuuskopio Hämeen ammattikorkeakoulun Onedriveen. Tietoja säilytetään vuoden ajan opinnäytetyön hyväksymisestä, minkä jälkeen ne poistetaan tietoturvallisesti.

Nykyisen järjestelmän toimittajan haastattelusta kerätyt muistiinpanot säilytetään toimeksiantajan omistamalla verkkolevyllä vähintään vuoden ajan opinnäytetyön hyväksymisestä.

Liite 2: Asiakaskysely

Kysely asiakasvirastoille henkilöstötutkimuksen avovastausten käsittelystä

1. Viraston tiedot *
 - 1.1. viraston / organisaation nimi
 - 1.2. henkilömäärä

2. Kuinka usein teette henkilöstölle tutkimuksia, joissa kerätään myös avoimia tekstivastauksia? *
 - a. Harvemmin kuin kerran vuodessa
 - b. 1–2 kertaa vuodessa
 - c. 3–5 kertaa vuodessa
 - d. Enemmän kuin 5 kertaa vuodessa

3. Kuinka paljon arvioitte työaikaa kuluvan avovastausten käsittelyyn yhden tutkimuksen osalta koko organisaatiossanne yhteensä?

4. Miten tällä hetkellä käsittelette henkilöstötutkimuksista saatavia avovastauksia?

5. Käytetäänkö avovastausten käsittelyyn tekoälyä tai muuta automatisointia? *
 - a. Kyllä
→ Miten tekoälyä tai muuta automatisointia käytetään avovastausten käsittelyyn?
 - b. Ei

6. Teettekö avovastauksista koosteita tai raportteja? *
 - a. Kyllä
→ Millaisia koosteita tai raportteja teette?
→ Teettekö koosteet itse vai ostatteko palveluna?
 - b. Ei

7. Miten hyödynnätte tutkimuksen avovastauksia?

Haastattelen muutamaa vastaajaa laajemmin henkilöstötutkimuksien avovastausten käsittelystä valtion organisaatioissa. Mikäli sinuun voi olla yhteydessä aiheeseen liittyen, jätähän alle yhteystietosi.

8. Sähköpostiosoite

Liite 3: Asiakashaastattelujen runko

Haastattelun kysymykset:

1. Kerro omin sanoin, miten lähdette käsittelemään avovastauksia, kun saatte ne esimerkiksi vuotuisen henkilöstötutkimuksen jälkeen.
2. Mikä toimii hyvin tällä hetkellä avointen vastausten käsittelyssä?
3. Mitä haasteita avointen vastausten käsittelyssä on?
4. Onko teillä käytössä tekoälyä muissa prosesseissa?
 - a. Missä?
 - b. Millainen käyttöönotto oli?
 - c. Miten henkilöstö on suhtautunut?
5. Millaisia koosteita voisitte ajatella käyttävänne, jos ne tuotettaisiin automaattisesti / tekoäly avusteisesti?
 - a. Sanapilvi – Eniten mainitut sanat
 - b. Teemat – Listaus eniten mainituista ennalta määritetyistä teemoista
 - c. Sentimentit – Avovastausten sävyt (positiivinen/negatiivinen) aiheittain tai kysymyksittäin
 - d. Muu, mikä?
6. Mitä hyötyjä tunnistatte avovastausten käsittelystä tekoälyllä?
7. Tunnistatteko haasteita / riskejä avovastausten käsittelystä tekoälyllä? Mitä?