



Vaasan yliopisto
UNIVERSITY OF VAASA

Roland Parts

Selitettävyyden ja vuorovaikutuksen kehittäminen tekoälypohjaisissa rekrytointijärjestelmissä

Työnhakijoiden näkökulma

Tekniikan ja innovaatiojohtamisen yksikkö
Kauppatieteiden Pro gradu -tutkielma
Teknisen viestinnän maisteriohjelma

Vaasa 2024

VAASAN YLIOPISTO**Tekniikan ja innovaatiojohtamisen yksikkö**

| | | | |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------|-------------------|-----|
| Tekijä: | Roland Parts | | |
| Tutkielman nimi: | Selitettävyyden ja vuorovaikutuksen kehittäminen tekoälypohjaisissa rekrytointijärjestelmissä: Työnhakijoiden näkökulma | | |
| Tutkinto: | Kauppätieteiden maisteri | | |
| Oppiaine: | Tekninen viestintä | | |
| Työn ohjaaja: | Tomi Pasanen | | |
| Valmistumisvuosi: | 2024 | Sivumäärä: | 117 |

TIIVISTELMÄ:

Tekoälypohjaisia rekrytointijärjestelmiä käytetään yhä enemmän työnhakijoiden seulonnessa ja arvioinnissa, vaikka niiden selitettävyyden ja vuorovaikutuksellisuus ovat jääneet vähälle huomiolle työnhakijoiden näkökulmasta. Tekoälypohjainen päätöksenteko voi vaikuttaa merkittävästi hakijoiden luottamukseen ja kokemukseen rekrytointiprosessista, mutta tähän mennessä tutkimus on keskittynyt pääasiassa järjestelmien tehokkuuteen. Selitettävyyden ja vuorovaikutuksellisuuden tarkastelu työnhakijoiden näkökulmasta on kuitenkin tarpeellista, sillä päätöksenteon läpinäkyvyys ja mahdollisuus vuorovaikutukseen voivat parantaa hakijoiden hyväksyntää ja sitoutumista rekrytointiprosessiin. Tämän tutkimuksen tavoitteena on kartoittaa, millaisia selitettävyyden- ja vuorovaikutusodotuksia työnhakijoilla on tekoälypohjaisissa rekrytointijärjestelmissä.

Tutkimus toteutettiin konstruktivisella tutkimusotteella ja aineisto kerättiin teemahaastattelulla, joissa haastateltiin työnhakijoita eri taustoista. Aineisto analysoitiin sisällönanalyysin avulla, mikä mahdollisti työnhakijoiden kokemusten ja näkökulmien syvällisen tarkastelun. Tutkimuksen tuloksena muodostettiin ohjeistus, joka käsittelee selitettävyyden ja vuorovaikutuksellisuuden kehityssuuntia tekoälypohjaisissa rekrytointijärjestelmissä. Ohjeistuksessa keskitytään selitysten selkeyteen, vuorovaikutusmahdollisuuksien lisäämiseen, kuten lisätietojen kysymiseen, sekä tekoälyn ja ihmisten välisten roolien selkeyttämiseen rekrytointiprosessin eri vaiheissa, kuten hakemisen, seulonnan ja arvioinnin yhteydessä. Ohjeistus sisältää konkreettisia suosituksia siitä, miten selitettävyyden voidaan huomioida järjestelmän käyttöliittymän suunnittelussa ja miten vuorovaikutuksellisuutta voidaan lisätä esimerkiksi tarjoamalla reaaliaikaista palautetta hakijoille heidän edistymisestään prosessin aikana.

Ohjeistuksen tarkoitus on tarjota rekrytointin kehittäjille kehityssuuntia, joiden avulla he voivat vastata työnhakijoiden toiveisiin selitettävyydestä ja vuorovaikutuksesta ja siten edistää hakijakokemusta tekoälypohjaisessa rekrytointiprosessissa. Jatkotutkimuksena voitaisiin arvioida ohjeistuksen käytettävyyttä ja vaikutuksia käytännön rekrytointitilanteissa sekä kartoittaa tarkemmin erilaisten työnhakijaryhmien, kuten eri ikäryhmien, näkemyksiä. Ohjeistuksen vaikutusta voitaisiin lisäksi arvioida käytännössä esimerkiksi pilottiohjelmien avulla, joissa tekoälypohjaisia rekrytointijärjestelmiä muokataan ohjeistuksen mukaisesti ja tarkastellaan, miten muutokset vaikuttavat hakijakokemukseen ja luottamukseen rekrytointiprosessiin.

AVAINSANAT: Selittävä tekoäly, läpinäkyvyys, rekrytointijärjestelmät, luotettavuus

Sisällys

| | | |
|-------|---------------------------------------------------------|----|
| 1 | Johdanto | 7 |
| 1.1 | Tutkimuksen tavoite ja menetelmä | 9 |
| 1.2 | Tutkielman rakenne | 9 |
| 2 | Tekoäly rekrytoinnissa | 10 |
| 2.1 | Tekoälyn määrittely ja lähestyminen | 10 |
| 2.2 | Tekoälyllä tehokkuutta rekrytointiprosessiin | 12 |
| 2.3 | Tekoälypohjaiset rekrytointijärjestelmät | 17 |
| 2.3.1 | Ennakoiva analytiikka | 18 |
| 2.3.2 | Algoritminen päätöksenteko | 20 |
| 2.3.3 | Keskustelevat käyttöliittymät | 21 |
| 2.4 | Tekoälyn vaikutuksia työnhakijan asemaan rekrytoinnissa | 23 |
| 3 | Selittävä tekoäly | 28 |
| 3.1 | Läpinäkyvyyden merkitys | 29 |
| 3.1.1 | Läpinäkyvyyden tasot tekoälyjärjestelmissä | 30 |
| 3.1.2 | Sidosryhmien intressit läpinäkyvyyden tasoilla | 33 |
| 3.1.3 | Selittävän tiedon ymmärtäminen | 37 |
| 3.2 | Selitettävyydestietoinen suunnittelu | 40 |
| 3.2.1 | Selitettävyydestavoitteet | 42 |
| 3.2.2 | Vaatimusten analysointi ja määrittely | 47 |
| 3.2.3 | Multimodaalinen vuorovaikutussuunnittelu | 49 |
| 3.2.4 | Vuorovaikutukselliset selitettävyydelähestymistavat | 52 |
| 4 | Tutkimusmenetelmä- ja aineisto | 55 |
| 4.1 | Tutkimusmenetelmä | 55 |
| 4.2 | Haastateltavat | 56 |
| 4.3 | Aineisto | 57 |
| 4.4 | Aineiston analyysi | 59 |
| 5 | Tutkimuksen tulokset | 62 |

| | |
|-------------------------------------------------------------------------------------------|-----|
| 5.1 Työnhakijoiden odotukset tekoälypohjaisten rekrytointijärjestelmien selitettävyydestä | 62 |
| 5.1.1 Tulosperusteisen selitettävyyden määrittäminen | 64 |
| 5.1.2 Prosessiperusteisen selitettävyyden määrittäminen | 67 |
| 5.2 Tekoälyjärjestelmien vuorovaikutuksellisuuden merkitys rekrytointiprosessissa | 68 |
| 5.2.1 Reaaliaikaisen informaation merkitys | 70 |
| 5.2.2 Mahdollisuus kysyä lisätietoja ja tarkennuksia | 71 |
| 5.2.3 Tekoälyn ja ihmisen roolien selkeys | 73 |
| 6 Toteutus | 75 |
| 6.1 Motivaatio ja vaatimukset | 75 |
| 6.2 Suunnittelu | 77 |
| 6.3 Ohjeistuksen toteutus | 84 |
| 6.4 Arviointi | 86 |
| 7 Johtopäätökset ja pohdinta | 90 |
| Lähteet | 93 |
| Liitteet | 105 |
| Liite 1. Haastattelurunko | 105 |
| Liite 2. Selitettävyyden ja vuorovaikutuksellisuuden konseptuaalinen kehittäminen | 108 |

Kuvat

| | |
|-----------------------------------------------------------------------------------------------------------------------------|----|
| Kuva 1. keskustelevien käyttöliittymien toimintoja ja rooleja rekrytointikontekstissa (Koivunen ja muut, 2022, s. 509–510). | 23 |
|-----------------------------------------------------------------------------------------------------------------------------|----|

Kuviot

| | |
|------------------------------------------------------------------------------------------------|----|
| Kuvio 1. Tekoälyn lähestymistavat tieteessä (Russell ja muut, 2022, s. 19). | 12 |
| Kuvio 2. Tehostavat roolit rekrytointiprosessissa (Nikolaou, 2021; Pillai ja muut, 2022). | 16 |
| Kuvio 3. Tekoälypohjaisten rekrytointijärjestelmien kehityssuuntia (Albassam, 2023, s. 65). | 18 |
| Kuvio 4. Päätöspuumalli (Zhang, 2022, s. 15 689). | 21 |
| Kuvio 5. Läpinäkyvyyden tasot tekoälyjärjestelmissä (Haresamudram ja muut, 2023, s. 95). | 31 |
| Kuvio 6. Sidosryhmien luokitus läpinäkyvyyden tasoilla (Barredo Arrieta ja muut, 2020, s. 84). | 34 |
| Kuvio 7. Selitystyyppit (Leslie & Briggs, 2021, s. 12). | 40 |
| Kuvio 8. Suunnittelulähestymistapojen hierarkia (Benyon, 2019; Bingley ja muut, 2023). | 42 |
| Kuvio 9. DoReMi-suunnittelumalli (Schoonderwoerd ja muut, 2021, s. 2). | 48 |
| Kuvio 10. Ennusteen ja selittämisen tuottamisen prosessi (Clement ja muut, 2023, s. 81). | 50 |
| Kuvio 11. Läpinäkyvien ja jälkiselitettävien mallien prosessi (Vilone & Longo, 2021, s. 616). | 51 |

Taulukot

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Taulukko 1. Vaikutuksista työntekijän asemaan hyötyjen ja riskien kannalta (Black & van Esch, 2020; Hunkenschroer & Luetge, 2022; Lookadoo & Moore, 2024). | 27 |
| Taulukko 2. Vaatimukset selitettävyydelle. | 76 |

Taulukko 3. Vuorovaikutukselliset vaatimukset.

77

1 Johdanto

Neljännän teollisen vallankumouksen mukanaan tuomat teknologiat, kuten tekoäly ja automaatio, ovat muuttaneet merkittävästi organisaatioiden sähköisiä rekrytointiprosesseja (Nikolaou, 2021, s. 2). Pillai ja muut (2022, s. 709–710) toteavat, että nykyisin teknologioiden avulla voidaan suorittaa monimutkaisia tehtäviä, kuten hakijoiden houkuttelua, arviointi ja valintaa, älykkäiden tietoteknisten järjestelmien avulla, jolloin toiminta jäljittelee rekrytointiammattilaisten päätöksentekoa. Tutkimukset osoittavat, että tekoälypohjaisia rekrytointijärjestelmiä hyödyntävät jo yli 90 % organisaatioista Yhdysvalloissa, Isossa-Britanniassa ja Saksassa, mikä viittaa merkittävään muutokseen henkilöstöhallinnoissa (Lookadoo & Moore, 2024, s. 3–4). Duongin & Thin (2022, s. 67) mukaan tekoälyn käyttö vaikuttaa erityisesti rekrytointiprosessien tehokkuuteen, mutta myös organisaatioiden ja hakijoiden välisen viestinnän kehittämiseen sillä tekoäly voi automatisoida viestintää ja parantaa tiedonkulkua rekrytointiprosessin eri vaiheissa.

Tekoälypohjaiset rekrytointijärjestelmät, erityisesti algoritmisen päätöksenteon, ennakoivan analytiikan ja keskustelevien käyttöliittymien muodossa, ovat lisänneet rekrytoinnin tehokkuutta, tarkkuutta ja oikeudenmukaisuutta sekä vähentäneet inhimillisten ennakkoluulojen vaikutusta (Albassam, 2023, s. 6–7; Hunkenschroer & Luetge, 2022, s. 997). Hunkenschroerin & Luetgen (2022, s. 997) mukaan teknologisen muutoksen keskellä on kuitenkin herännyt huolta tekoälyn käytön läpinäkyvyydestä ja selitettävyydestä, sillä työnhakijoiden näkökulmasta on edelleen epäselvää, millä tiedoilla heitä arvioidaan ja kuinka nämä tiedot vaikuttavat valintapäätöksiin. Läpinäkyvyyden puute rekrytointijärjestelmissä voi heikentää työnhakijoiden luottamusta prosessiin, sillä monimutkaisia järjestelmiä pidetään usein vaikeasti selitettävissä olevina ja lisäksi on pelko, että valintamenettelyjä saatettaisiin manipuloida (Van Iddekinge ja muut, 2023, s. 671).

Selittävän tekoälyn lähestymistavat tarjoavat potentiaalisia ratkaisuja tekoälyn ja ihmisen välisen vuorovaikutuksellisen läpinäkyvyyden kehittämiseksi, mikä korostaa myös työnhakijoiden oikeudenmukaisia mahdollisuuksia ymmärtää valintaprosessia (Haresamudram ja muut, 2023, s. 97–98). Haresamudram ja muut (2023) osoittavat, että keskustelu tekoälyjärjestelmien läpinäkyvyydestä on kehittynyt pitkään ikään kuin vuorovaikutuksen elementit eivät vaikuttaisi siihen. Heidän mukaansa päätösten selitettävyyttä sekä tekoälyjärjestelmien ja ihmisen välistä vuorovaikutusta läpinäkyvyyden moniulotteisuuden kannalta on tutkittu toistaiseksi vähän erilaisissa konteksteissa. Langer & König (2023) ovat tutkineet henkilöstöhallinta- ja rekrytointijärjestelmien algoritmisten päätösten läpinäkymättömyyden mahdollisia negatiivisia vaikutuksia eri käyttäjäryhmien näkökulmista. Lisäksi Langer ja muut (2021) tarkastelevat keskeisiä käsitteitä ja suhteita, joita tulisi huomioida, kun valitaan, arvioidaan, säädetään tai kehitetään selitettävyytlähestymistapoja, joiden tavoitteena on tyydyttää järjestelmän käyttäjien toiveita selitettävyydestä. Chromik & Butz (2021) ovat tutkineet vuorovaikutussuunnittelua selittävän tekoälyn periaatteiden näkökulmasta ja muodostivat vuorovaikutuskonseptit selittävän tekoälyn periaatteiden mukaisiksi. Leslie & Briggs (2021) esittelevät tapaustutkimuksessaan ohjeistuksen, jota voidaan hyödyntää tekoälypohjaisten rekrytointijärjestelmien selitystietoisessa suunnittelussa.

Selittävän tekoälyn lähestymistapoja rekrytointiprosessien läpinäkyvyyden ja vuorovaikutuksen parantamiseksi työnhakijoiden näkökulmasta on tutkittu toistaiseksi vain vähän. Tämä tutkielma keskittyy selittävän tekoälyn vuorovaikutuksellisen läpinäkyvyyden suunnitteluun ja tekoälyperusteisten rekrytointijärjestelmien selitettävyyden kehittämiseen työnhakijoiden tarpeet huomioiden. Tutkimuksessa tarkastellaan työnhakijoita kohderyhmänä ja analysoidaan heidän asemaansa ja odotuksiaan rekrytointiprosessissa. Empiirisen tutkimuksen pohjalta muodostetaan ohjeistus, joka esittelee keskeiset selitettävyytlähestymistavat ja vuorovaikutuksellisuuden ratkaisut, joilla voidaan tukea työnhakijoiden ymmärrystä ja osallisuutta tekoälypohjaisissa rekrytointijärjestelmissä.

1.1 Tutkimuksen tavoite ja menetelmä

Tämän tutkimuksen tavoitteena on selvittää, miten tekoälypohjaisten rekrytointijärjestelmien vuorovaikutuksellista läpinäkyvyyttä voidaan kehittää. Tutkimusmenetelmänä käytetään konstruktivistista tutkimusotetta. Tutkimuskysymykset ovat:

”Millaisia selitettävyyslähestymistapoja työnhakijat odottavat tekoälypohjaisilta rekrytointijärjestelmiltä?”

”Miten tekoälypohjaisten rekrytointijärjestelmien tekemät päätökset ja käsittelemät tiedot tulisi esittää työnhakijoille rekrytointiprosessin aikana?”

1.2 Tutkielman rakenne

Tutkielma koostuu seitsemästä kappaleesta. Johdannossa esitetään tutkimuksen aihe, tavoitteet, tarkoitus, menetelmä ja rakenne. Teoria on jaettu kahteen pääkappaleeseen, jotka ovat tekoäly rekrytointijärjestelmissä ja selittävä tekoäly. Tutkimusmenetelmä kappaleessa käsitellään, miten tutkimus suunniteltiin ja toteutettiin. Tutkimustulokset esitetään viidennessä kappaleessa. Toteutuskappaleessa esitetään, miten konstruktiio suunniteltiin ja toteutettiin. Lopuksi esitetään tutkimuksesta tehdyt johtopäätökset ja pohdinta.

2 Tekoäly rekrytoinnissa

Neljännän teollisen vallankumouksen myötä useat eri alat, mukaan lukien henkilöstöhallinto ja sen käyttämät rekrytointijärjestelmät, ovat kokeneet merkittäviä muutoksia (Norman, 2022, s. 37). Pillai ja muiden (2022, s. 709–710) mukaan tekoäly on keskeisin teknologia rekrytointijärjestelmien uudistuksessa, sillä sen avulla organisaatiot voivat automatisoida tiedonkäsittelyä ja päätöksentekoa tehokkaasti. Myös Black & van Esch (2020, s. 218–219) korostavat automaation merkitystä erityisesti rekrytointiprosessin neljässä keskeisessä vaiheessa, joita ovat markkinointi, seulonta, arviointi ja koordinointi. Henkilöstöhallinnon alalla tekoälypohjainen prosessiautomaatio tarkoittaa, että ohjelmistot voivat tehdä ja toteuttaa päätöksiä algoritmisesti usein ilman merkittävää ihmisen osallistumista (Johnson ja muut, 2021, s. 42).

Tässä kappaleessa ja sen alaluvuissa tarkastellaan tekoälyn roolia sekä merkitystä rekrytoinnin eri vaiheissa. Ensin määritellään tekoäly ja sen lähestymistavat ja analysoidaan, miten tekoäly voi tehostaa rekrytointiprosessia. Tämän jälkeen käsitellään tekoälypohjaisia rekrytointijärjestelmiä ennakoivan analytiikan, algoritmisen päätöksenteon ja keskustelevien käyttöliittymien näkökulmista. Lopuksi pohditaan tekoälyn vaikutuksia työnhakijoiden asemaan ja heidän kohtaamiinsa ongelmiin rekrytointiprosessissa. Tämä kokonaisuus tarjoaa kattavan kuvan tekoälyn monipuolisista rooleista ja vaikutuksista rekrytointikontekstissa.

2.1 Tekoälyn määrittely ja lähestyminen

Tekoäly viittaa tietokoneiden ja ohjelmistojen kykyyn suorittaa tehtäviä, jotka tavallisesti edellyttävät inhimillistä älykkyyttä (Osaba, 2021, s. 3). Campesaton (2020, s. 5–6) mukaan tekoälyn soveltamia tehtäviä voivat olla päätöksenteko, oppiminen, ongelmanratkaisu ja kielen ymmärtäminen, ja näiden tehtävien moninaisuus ulottuu yksinkertaisista automaatiotehtävistä monimutkaisiin päätöksentekoprosesseihin.

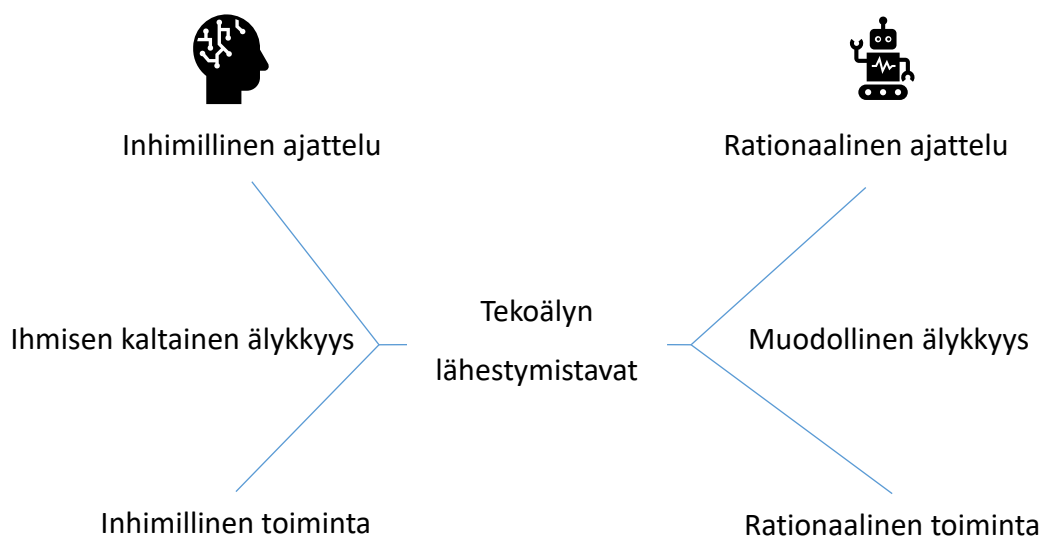
Keskeisenä osana tekoälyn toimintaa on sen kyky hyödyntää suuria määriä dataa oppimisprosesseissaan (Russell ja muut, 2022, s. 44).

Tekoälyjärjestelmät analysoivat ja prosessoivat dataa tunnistukseen olennaisia tietoja ja kaavoja, mikä mahdollistaa niille päätelmien tekemisen, oppimisen kokemuksista ja mukautumisen uusiin tilanteisiin (Russell ja muut, 2022, s. 43). Russell ja muut (2022, s. 44) korostavat, että tekoälyn keskeinen tarkoitus on luoda lisäarvoa kerätystä datasta. Osaba (2021, s. 2) toteaa, että tekoälyn viitekehykseen liittyy monia alasektoreita, joista erityisesti koneoppiminen, syväoppiminen ja optimointi korostuvat niiden datakeskeisten menetelmien vuoksi, sillä ne perustuvat tietojenkäsittelyn ja tilastotieteen käsitteisiin. Tekoälyn keskeisiä osa-alueita ovat myös esimerkiksi luonnollisen kielen käsittely, tiedon mallintaminen, robotiikka ja automaattinen päättely (Campesato, 2020, s. 18).

Tekoäly on monimuotoinen tieteenala, jonka tavoitteena on kehittää älykkäitä järjestelmiä reaali maailman ongelmien ratkaisemiseen (Osaba, 2021, s. 2). Russell ja muut (2022, s. 19) esittävät, että tekoälytutkimusta voidaan tarkastella ihmisten kaltaisuuden ja rationaalisuuden sekä ajatteluun ja käyttäytymiseen perustuvien mallien kautta. Neljä keskeistä lähestymistapaa, joita tekoälytutkimuksessa käytetään, ovat inhimillinen ajattelu, rationaalinen ajattelu, inhimillinen toiminta ja rationaalinen toiminta (Russell ja muut, 2022, s. 20).

Vold (2024, s. 231–232) määrittelee inhimillisen ajattelun lähestymistavaksi, jossa keskitytään mallintamaan ihmisen kognitiivisia toimintoja, kuten päättelyä ja muistia. Rationaalinen ajattelu puolestaan perustuu matemaattisiin ja teknisiin periaatteisiin ja pyrkii loogisiin päätöksiin. Russell ja muut (2022, s. 19) kuvaavat inhimillistä toimintaa ja ajattelua lähestymistapana, jossa järjestelmät toimivat ympäristössä ihmisten tavoin. Kun taas Fanin ja muiden (2022, s. 2–3) mukaan, että rationaalinen toiminta ja ajattelu tähtää optimaaliseen toimintaan tietyssä ympäristössä. Tekoälytutkimuksen lähestymistapoja ohjaa kaksijakoinen pyrkimys, joka jakautuu ihmisen kaltaisen

älykkyyden tavoitteluun ja muodollisen älykkyyden määrittelyyn (Russell ja muut, 2022, s. 19). Alla oleva kuvio 1 havainnollistaa tekoälytutkimuksen lähestymistapoja. Tässä tutkielmassa tarkastellaan erityisesti ihmisen kaltaista tekoälyä ja pyritään empiirisen tiedon avulla selvittämään työnhakijoiden odotuksia selitettävyydestä ja vuorovaikutuksesta tekoälypohjaisissa rekryointijärjestelmissä.



Kuvio 1. Tekoälyn lähestymistavat tieteessä (Russell ja muut, 2022, s. 19).

2.2 Tekoälyllä tehokkuutta rekryointiprosessiin

Nikolaou (2021, s. 2) määrittelee rekrytoinnin laajaksi käsitteeksi, joka kattaa monia eri elementtejä ja koostuu sarjasta prosesseja, järjestelmiä ja strategioita. Standardisoitu rekryointiprosessi sisältää yleensä kuusi päävaihetta, jotka ovat tarve, houkuttelu, seuranta, arviointi, valinta ja palkkaus (Lee, 2007, s. 83). Lee (2007, s. 83) kuvaa, kuinka tarvevaiheessa organisaatioissa tunnistetaan työtarve, jonka pohjalta luodaan kandidaattiprofiili avoinna olevan tehtävän vaatimusten mukaisesti. Tämä profiili

Julkaistaan digitaalisissa kanavissa työpaikkailmoituksena organisaation verkkosivuilla ja työpaikkasivustoilla. Tekoälyä voidaan hyödyntää tarvemäärittelyn analysoinnissa ja mallintamisessa, jolloin kandidaattiprofiilien luomisesta tulee tarkempaa ja osuvampaa (Pillai ja muut, 2022, s. 710). Stone ja muut (2015, s. 217) toteavat, että työpaikkailmoituksen jälkeen siirrytään houkutteluvaiheeseen, jossa painopiste on tehokkaassa markkinoinnissa ja viestinnässä potentiaalisille hakijoille.

Houkutteluvaihe on kehittynyt teknologian myötä menetelmiin, joissa hyödynnetään erilaisia digitaalisia työkaluja, kuten älykkäitä chatbotteja, joiden avulla pyritään tunnistamaan hakijoiden soveltuvuutta ja houkuttelemaan työtehtävään hakemista (Nikolaou, 2021, s. 2). Nikolaou (2021, s. 2) toteaa, että digitaalisia työkaluja hyödyntämällä pyritään luomaan kiinnostavaa sisältöä, kuten verkkopohjaisia työpaikkailmoituksia sekä sosiaalisen median kampanjoita, joiden avulla voidaan tavoittaa laajempi ja monimuotoisempi hakijajoukko. Stone ja muut (2015, s. 217) korostavat, että vahva työnantajabrändi yhdessä tehokkaiden houkuttelutoimenpiteiden kanssa on keskeinen tekijä laadukkaan hakijapoolin saamisessa, mikä puolestaan lisää mahdollisuuksia löytää oikea osaaja työtehtäviin. Houkutteluvaiheen aikana tai sen jälkeen alkaa yleensä rekrytointiprosessin hallinnollinen seurantavaihe (Lee, 2007, s. 84).

Lee (2007, s. 84) kuvaa, että seurantavaiheeseen kuuluu työhakemusten kerääminen, hakijoiden ansioluetteloiden tallentaminen osaksi ehdokasprofiileja ja hakijoiden tilanteen seuraaminen rekrytointiprosessissa. Tässä vaiheessa standardimenetelmä on, että työnhakija voi luoda profiilin organisaation verkkosivuilla ja täyttää verkkolomakkeen. Lomakkeeseen täytetään tarvittavat tiedot ja liitetään pyydetty tiedostot, kuten ansioluettelo ja valokuva (Laumer ja muut, 2015, s. 432). Lee (2007, s. 85) toteaa, että seurantavaiheen merkitys rekrytointiprosessissa korostuu erityisesti hakijoiden tietojen systemaattisessa hallinnassa ja hakuprosessin läpinäkyvyydessä. Tässä vaiheessa tekoälyjärjestelmät tehostavat hakijoiden tietojen analysointia ja profilointia, mikä nopeuttaa hakemusten käsittelyä ja parantaa ehdokkaiden ja organisaation välistä yhteensovittamista (Pillai ja muut, 2022, s. 710).

Arviointivaihe sisältää yleensä kaksiosaisen prosessin, jossa ensin hakijan ansioluettelo ja hakemuskirje seulotaan avainsanojen avulla, minkä jälkeen hakijaa arvioidaan syvällisemmin haastattelujen ja testien avulla (Lee, 2007, s. 84). Nikolaou (2021, s. 3) toteaa, että on yhä yleisempää tarkastella hakijoiden sosiaalisen median profiileja, kuten LinkedIniä ja Facebookia, heidän persoonallisuutensa arvioimiseksi. Faliagka ja muut (2012) esittävät lähestymistavan, jossa hakijat analysoidaan persoonallisuuden perusteella ja sijoitetaan paremmuusjärjestykseen heidän kompetenssiensa perusteella. Tekoälyä käytetään tehostamaan arviointivaihetta esimerkiksi automaattiseen analysointiin ja hakijoiden vertailuun, mikä mahdollistaa nopeamman, syvemmän ja tarkemman päätöksenteon (Pillai ja muut, 2022, s. 710). Nikolaou (2021, s. 3) mukaan tekoälyn avulla voidaan parantaa myös työnhakijoiden ja organisaatioiden välistä viestintää, mikä voi ilmetä automaattisena tiedottamisena, esimerkiksi hakijoiden työnhakuun liittyvistä suorituksista ja pisteytyksistä.

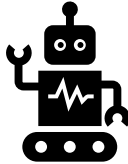
Viestinnän merkitys korostuu erityisesti arviointi- ja valintavaiheessa, sillä nämä vaiheet sisältävät työnhaun kannalta merkityksellisiä ja sensitiivisiä henkilötietoja (Holm, 2012, s. 251–252). Lee (2007, s. 84) toteaa, että arviointivaiheen tavoitteena on valita potentiaalisimmat ehdokkaat, joita voidaan tarvittaessa haastatella uudelleen ennen lopullista valintaa ja palkkausta. Digitaaliset työkalut mahdollistavat myös videohaastattelujen toteuttamisen, jolloin hakijat voivat suorittaa haastattelut ajasta ja paikasta riippumatta (Nikolaou, 2021, s. 3). Toisaalta Langerin ja muiden (2017, s. 378) mukaan merkittävä osa hakijoista kokee digitaaliset haastattelut kuitenkin pelottavammiksi ja vähemmän henkilökohtaisiksi perinteisiin haastatteluihin verrattuna. Tekoälyn vaikutus videohaastatteluihin ilmenee automaatiassa, jossa esimerkiksi tekoälyagentit haastattelevat ja analysoivat hakijoita ihmisten sijasta (Pillai ja muut, 2022, s. 710).

Yksinkertaisimmillaan rekrytoinnin tavoitteena on palkata oikeat henkilöt oikeisiin tehtäviin hyödyntämällä tehokkaita prosesseja, järjestelmiä ja strategioita (Lee, 2007, s.

85). Nikolaou (2021, s. 3) toteaa, että teknologiakehitys on ohjannut rekrytointiprosessien, -järjestelmien ja -strategioiden kehittämistä erityisesti houkutteluun, arviointiin ja viestintään liittyviin osa-alueisiin. Hunkenschroer & Luetge (2022, s. 998) kuitenkin huomauttavat, että rekrytointikehityksessä, kuten tekoälyperusteisissa rekrytointijärjestelmissä, painopiste on usein ollut organisaatioiden tehokkuuden parantamisessa sen sijaan, että keskittyisi kandidaattikokemuksen kehittämiseen. Tämä on osittain ymmärrettävää, sillä organisaatiot voivat saavuttaa kustannussäästöjä, jotka ylittävät jopa yli 50 prosenttia ei-tekoälypohjaisiin rekrytointijärjestelmiin verrattuna (Pillai ja muut, 2022, s. 709). Hunkenschroer & Luetge (2022, s. 1001) korostavat, että tekoälyperusteisissa rekrytointijärjestelmissä tulisi tulevaisuudessa kiinnittää enemmän huomiota kandidaattikokemuksen parantamiseen, sillä esimerkiksi järjestelmien automaattisen päätöksenteon läpinäkyvyys on tärkeää eettisyyden kannalta. Alla oleva kuvio 2 havainnollistaa tekoälyn roolia rekrytointiprosessien tehostavana työkaluna.

Tarvevaihe

- Tukea tarvemäärittelyn analysointiin ja mallintamiseen



**Tekoälyn tehostavat roolit
rekrytointiprosesseissa**

Valintavaihe

- Videohaastattelut ja tekoälyagentit: haastattelujen järjestäminen ja hakijoiden analysointi

Houkutteluvaihe

- Chatbotit: Hakijoiden soveltuvuuden tunnistamiseen ja hakijoiden houkuttelemiseen tehtävään
- Sisällöntuotanto: Kiinnostavan sisällön luomiseen

Seurantavaihe

- Hakemusten ja ansioluetteloiden hallinta: Tehostavat hakijoiden tietojen analysointia ja profilointia
- Hakemusten käsittelyn nopeutus ja yhteensovittaminen: Parantaa ehdokkaiden ja organisaatioiden välistä yhteensovittamista analysoimalla tietoja

Arviointivaihe

- Avainsanojen analysointi: ansioluetteloiden ja hakemuskirjeiden avainsanaseulonta
- Hakijoiden sosiaalisen median profiilit: mahdollistaa sosiaalisen median profiilien analysoinnin persoonallisuuden arvioimiseksi
- Automatisoitu analyysi ja vertailu: syvempää hakijoiden vertailua ja arviointi päätöksenteon tueksi
- Viestintä hakijoiden kanssa: automaattista tiedottamista rekrytoinnista

Kuvio 2. Tehostavat roolit rekrytointiprosessissa (Nikolaou, 2021; Pillai ja muut, 2022).

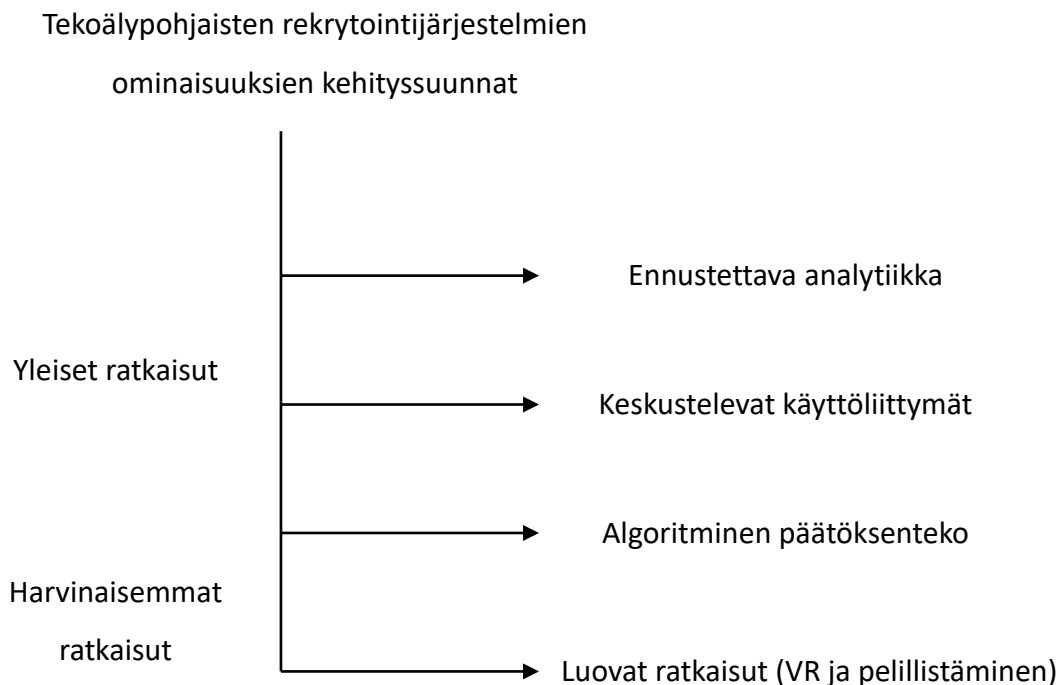
2.3 Tekoälypohjaiset rekrytointijärjestelmät

Rekrytointijärjestelmän tekninen arkkitehtuuri on yleensä laaja kokonaisuus, joka koostuu erilaisista alijärjestelmistä, joita käytetään rekrytointiprosessiin liittyvien tehtävien suorittamiseen (Lee, 2007, s. 82). Lee (2007, s. 83–84) toteaa, että alijärjestelmiä voi olla useita, ja jokaisella niistä on oma erityinen tehtävänsä, kuten viestintä, hallinnointi, arviointi ja analysointi. Rekrytointijärjestelmän alijärjestelminä voivat olla esimerkiksi työtarpeenhallintajärjestelmä, rekrytoinnin suorituskykyanalyysijärjestelmä, arviointijärjestelmät, viestintäjärjestelmät sekä esikarsintahallintajärjestelmä (Laumer ja muut, 2015, s. 433).

Rekrytointijärjestelmän alijärjestelmät voidaan nähdä rakenteistettuina tietoteknisinä komponentteina, joilla on määriteltyjä ominaisuuksia ja toiminnallisuuksia rekrytointitehtävien suorittamiseen (Laumer ja muut, 2015, s. 434). Laumer ja muut (2015, s. 432–433) kuvaavat, että rekrytointiprosessit jäsennellyiksi liiketoiminnan ydinprosesseiksi, joita virtaviivaistetaan tietoteknisten komponenttien avulla. Rekrytointiprosessissa tietotekninen kehitys on erityisesti keskittynyt houkuttelu-, arviointi- ja valintavaiheisiin sekä niihin liittyviin analysointityökaluihin (Nikolaou, 2021, s. 2–3). Pillai ja muut (2022, s. 710) korostavat, että tekoäly tukee näiden komponenttien kehittämistä parantamalla prosessien tehokkuutta ja tarkkuutta, mikä vaikuttaa myönteisesti rekrytointituloksiin.

Tekoälypohjaisissa rekrytointijärjestelmissä korostuvat erityisesti, ennustettava analytiikka, algoritminen päätöksenteko ja keskustelevat käyttöliittymät (Albassam, 2023, s. 65). Pillai ja muut (2022, s. 711) toteavat, että tekoäly mahdollistaa myös luovien rekrytointijärjestelmien kehittämisen, esimerkiksi virtuaalisen todellisuuden ja pelillistämisen hyödyntämisen simuloituissa työtehtävien arvioinneissa. Luovat rekrytointijärjestelmäratkaisut ovat kuitenkin harvinaisempia, sillä tekoälypohjaisten rekrytointijärjestelmien kehittäminen on keskittynyt ensisijaisesti datan hyödyntämiseen algoritmisessa päätöksenteossa, rekrytointitietojen data-analyyseissä

ja vuorovaikutuksellisissa keskustelemissa käyttöliittymissä (Woods ja muut, 2020). Alla oleva kuvio 3 havainnollistaa tekoälypohjaisten rekrytointijärjestelmien kehityssuuntia.



Kuvio 3. Tekoälypohjaisten rekrytointijärjestelmien kehityssuuntia (Albassam, 2023, s. 65).

2.3.1 Ennakoiva analytiikka

Albassam (2023, s. 6–7) toteaa, että ennakoiva analytiikka hyödyntää Big Dataa, tiedonlouhintaa ja koneoppimisalgoritmeja mallien tunnistamiseksi ja tulevien lopputulosten ennustamiseksi. Big Data ylittää perinteisten tietokantajärjestelmien käsittelykyvyn ja eroaa perinteisistä tietorakenteista (Zhang, 2022, s. 15 685). Zhangin 2022 (s. 15 686) mukaan Big Datasta etsitään malleja, trendejä tai hyödyllistä tietoa tiedonlouhinnan keinoin. Tiedonlouhintamenetelmiä ovat esimerkiksi luokittelu, klusterointi, assosiaatiosääntöjen löytäminen ja poikkeamien tunnistaminen (Gao, 2021, s. 3).

Rekrytoinnin yhteydessä ennakoivaa analytiikkaa käytetään historiallisten tietojen analysoimiseen mallien tunnistamiseksi (Albassam, 2023, s. 14). Pillai ja muut (2022, s. 710–711) korostavat, että rekrytoinnissa kehitetään ennustemalleja tarkasti kohdistettuihin tarkoituksiin, kuten rekrytointimainontaan, kandidaattien arviointiin, heidän potentiaalisen suorituskäytönsä ennustamiseen ja rekrytointitarpeiden arviointiin. Ennakoivan analytiikan tekniikoilla pyritään saamaan tarkempaa ja laadukkaampaa tietoa, jotta rekrytointitehtävissä voidaan tehdä perusteltuja ratkaisuja (Albassam, 2023, s. 14).

Organisaatiot tunnistavat dataohjatun rekrytointistrategian hyödyt ja pyrkivät hyödyntämään sitä strategisessa päätöksenteossa, sillä se parantaa yrityksen suorituskäytönsä liiketoiminnallisesti (Conte & Siano, 2023, s. 622). Dahlbom ja muut (2020, s. 627) huomauttavat, että dataohjattu strategia voi tarjota merkittävää lisäarvoa rekrytoinnin päätöksenteolle ja viestinnälle, mutta onnistunut käyttöönotto vaatii konkreettisia merkittäviä kehitystoimenpiteitä ja investointeja. Käyttöönoton esteinä ovat sekä tekniset inhimilliset ongelmat, kuten heikko datan laatu ja henkilöstön riittämättömät analytiikkataidot (Dahlbom ja muut, 2020, s. 628).

Investoimalla henkilöstön koulutukseen ja parantamalla datan laatua, organisaatiot voivat saavuttaa merkittäviä etuja rekrytoinnissa (Dahlbom ja muut, 2020, s. 627). (Hunkenschroer & Luetge, 2022, s. 999) korostavat, että organisaatioiden tulisi pohtia myös eettisiä näkökulmia datan käytössä, jotta analytiikan hyödyntäminen olisi vastuullista ja läpinäkyvää. Rekrytointijärjestelmien tulisi esimerkiksi välttää hakijoiden syrjintää algoritmien käytössä varmistamalla, ettei analytiikka suosi tiettyjä hakijaryhmiä iän, sukupuolen tai etnisen taustan perusteella. Tämä on erityisen tärkeää Euroopassa, sillä Euroopan unionin parlamentti tiukentaa säännöksiä, jotka koskevat tekoälyn käyttöä ja sen eettisiä periaatteita (*AI Pact | Shaping Europe's Digital Future*, 2024).

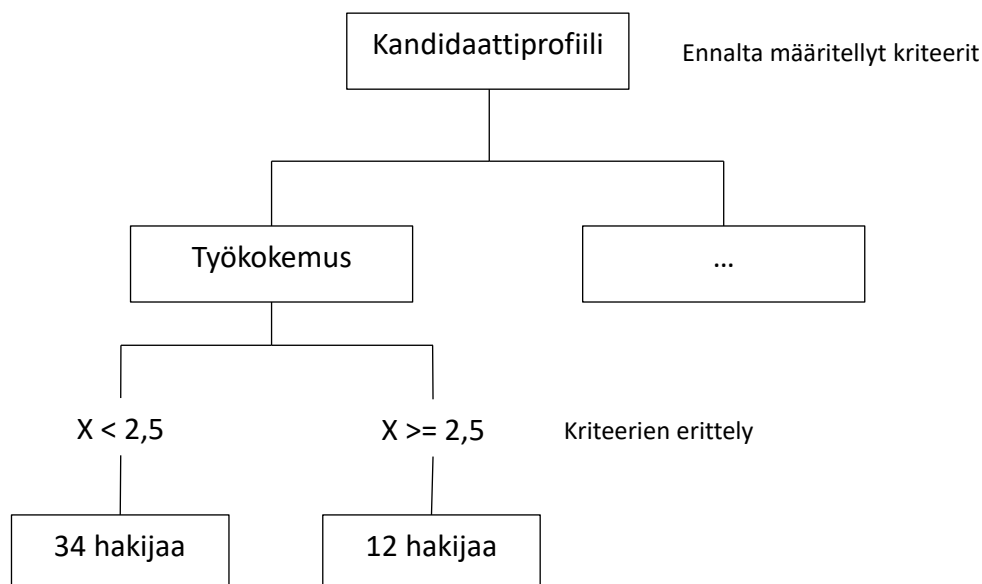
2.3.2 Algoritminen päätöksenteko

Algoritminen päätöksenteko tarkoittaa prosessia, jossa algoritmit tekevät päätöksiä automaattisesti ihmisen sijasta tai ovat siinä vain vähäisesti tukena (Köchling ja muut, 2023, s. 2110). Köchling ja muut (2023, s. 2110–2111) toteavat, että algoritmit käyttävät dataa ja data-analytiikan menetelmiä analysoidakseen tilanteita ja tehdäkseen päätöksiä määritettyjen matemaattisten periaatteiden, kuten sääntöjen, ehtojen tai ennustemallien, perusteella. Rekrytoinnissa algoritminen päätöksenteko vaikuttaa yleisemmin seulontaan, jossa tekoälyjärjestelmä arvioi työnhakijoiden ansioluetteloita, testituloksia ja muita rekrytointiin liittyviä tietoja ennustemallien avulla (Woods ja muut, 2020, s. 65–66).

Algoritmisella päätöksenteolla on tarkoitus parantaa rekrytointiprosessin nopeutta ja tehokkuutta, sillä algoritmit voivat analysoida suuria tietomääriä huomattavasti nopeammin kuin ihmiset (Albassam, 2023, s. 7). Pillai ja muiden (2022, s. 710) mukaan tekoäly mahdollistaa sen, että hakemuksia voidaan käsitellä ja arvioida lähes reaaliaikaisesti, mikä voi nopeuttaa rekrytointiprosessia. Lisäksi algoritmit tarjoavat objektiivisuutta päätöksenteossa, sillä ne voivat perustaa valintansa selkeästi määriteltyihin kriteereihin ilman inhimillisiä ennakkoluuloja tai tunteita, mikä voi johtaa tasapuolisempiin lopputuloksiin (Woods ja muut, 2020, s. 65).

Algoritmisessa päätöksenteossa on erilaisia vaihtoehtoja data-analyttisiin menetelmien suhteen, kuten esimerkiksi päätöspuumalli, joka on yksi suosituimmista menetelmä, sillä se tarjoaa selkeän ja helposti tulkittavan rakenteen päätöksentekoprosessille (Zhang, 2022, s. 15 689). Zhang (2022, s. 15 689–15 690) kuvaa, että päätöspuumalli jakaa tietojoukon hierarkkiseen rakenteeseen, jossa päätöksiä tehdään haarojen kautta. Tämä mahdollistaa selkeän visualisoinnin ja analyysin vaikutuksista sekä auttaa tunnistamaan tärkeimmät muuttujat ja niiden suhteet. Päätöspuumallit ovat erityisen käyttökelpoisia rekrytoinnin monimutkaisissa päätöksentekotilanteissa, sillä ne pystyvät luokittelemaan tietoja tarkasti ja erittelemään kriteerejä, kuten hakijan pätevyyttä, kokemusta ja

soveltuvuutta (Faliagka ja muut, 2012, s. 557). Alla oleva kuvio 4 havainnollistaa päätöspuumallia.



Kuvio 4. Päätöspuumalli (Zhang, 2022, s. 15 689).

2.3.3 Keskustelevat käyttöliittymät

Keskusteleva käyttöliittymä, kuten chattibotti, on ohjelmistoratkaisu, joka hyödyntää luonnollisen kielen käsittelyä kommunikoidakseen käyttäjien kanssa (Albassam, 2023, s. 13). Wiberg & Stolterman Bergqvist (2023, s. 2 287) toteavat, että OpenAI kehittämä ChatGPT on hyvä esimerkki tekoälypohjaisesta chattibotista, joka käyttää kone- ja syväoppimistekniikoita tuottaakseen geneerisesti ihmisen kaltaisia vastauksia tekstipohjaisiin keskusteluihin. ChatGPT-chatbotissa korostuvat korkea automaatio ja vuorovaikutus käyttäjän kanssa, sillä se vastaa kysymyksiin ja rakentaa keskustelua aiempien kehoitteiden ja vastausten pohjalta, mikä mahdollistaa jatkuvan vuoropuhelun (Wiberg & Stolterman Bergqvist, 2023, s. 2 287).

Geneeristen tekoälyjen vaikutus keskusteleviin käyttöliittymiin voi vaihdella merkittävästi eri konteksteissa, mutta rekrytoinnissa niiden käyttö on erityisen lupaavaa

työnhakijoiden ja työnantajien välisessä viestinnässä sekä vuorovaikutuksessa (Budhwar ja muut, 2023, s. 622). Koivusen ja muiden (2022, s. 512) mukaan chattibotit yhdessä tekoälyn kanssa mahdollistavat älykkäiden keskustelevien käyttöliittymien kehittämisen rekrytointitoimintojen sekä prosessien tueksi organisaatioissa, joissa erityisen hyödyllisenä koetaan tiedonkeruun ja viestinnän automatisointi. Nämä keskustelevat käyttöliittymät voivat parantaa rekrytointiprosessien ja -toimintojen koordinoimista sekä seurantaan ja tehostaa automaattisen viestinnän kautta tavoiteltujen kandidaattien houkuttelua (Black & van Esch, 2020, s. 222).

Rekrytointikontekstissa chatbotit toimivat tyypillisesti tehtävä- ja palvelulähtöisesti (Koivunen ja muut, 2022, s. 509). Koivunen ja muut (2022, s. 509) toteavat, että chattibotit voivat toteuttaa rekrytointitoimintoja itsenäisesti. Esimerkiksi houkuttelubotti pyrkii tavoittamaan potentiaalisia kandidaatteja automaattisella viestinnällä ja keräämään rekrytointitietoja, kuten ansioluetteloja. Organisaatiot voivat hyödyntää houkutteluchattibotteja laaja-alaisesti rekrytoinnissa, esimerkiksi arvioimalla hakijoiden tietoja, suorittamalla esivalintaa niiden perusteella ja suosittelemalla avoimia työtehtäviä hakijoiden profiilien mukaan (Pillai ja muut, 2022, s. 716).

Koivunen ja muut (2022, s. 510) analysoivat, että chattibotti-käyttöliittymiä voidaan potentiaalisesti käyttää myös työnhakijoille tarkoitetulla rekrytoinnin seuranjärjestelmänä. Tämä palveluorientoitunut lähestymistapa voisi olla tarpeellista, sillä moni työnhakija kokee saavan liian vähän tietoa rekrytointiprosessin aikana (Black & van Esch, 2020, s. 222). Black & van Esch (2020, s. 222) toteavat, että chattiboteilla voidaan parantaa ennakoivaa tiedottamista, jonka avulla hakijoille kerrotaan, missä vaiheessa he ovat rekrytointiprosessissa ja mitkä ovat seuraavat vaiheet. Tekoälypohjaisissa rekrytointijärjestelmissä läpinäkyvyys on erityisen tärkeää, koska se voi merkittävästi parantaa hakijoiden kokemusta ja vähentää epävarmuutta rekrytointiprosessin aikana (Hunkenschroer & Luetge, 2022, s. 997). Alla oleva kuvassa 1 on esitelty keskustelevien käyttöliittymien toimintoja ja rooleja rekrytointikontekstissa.



Keskustelevien käyttöliittymien toimintoja ja rooleja



Houkutteleva ja viestintä:

Chatbotit tavoittavat ja houkuttelevat potentiaalisia työnhakijoita, esimerkiksi automaattisilla viesteillä ja tiedonkeruulla.



Esivalinta ja arviointi:

Chatbot voi suorittaa alustavaa tietojen keräämistä ja arviointia, esimerkiksi ansioluetteloiden analysointia.



Rekrytoinnin seurantajärjestelmä:

Chatbotit voivat toimia työnhakijoiden seurantajärjestelmänä, jossa hakijat voivat tarkistaa prosessin etenemistä ja seuraavia vaiheita.



Tiedonkeruu ja prosessin seuranta:

Chatbot mahdollistaa työnhakijoiden tietojen keräämistä ja antaa hakijoille tietoa rekrytointiprosessin vaiheista.

Kuva 1. keskustelevien käyttöliittymien toimintoja ja rooleja rekrytointikontekstissa (Koivunen ja muut, 2022, s. 509–510).

2.4 Tekoälyn vaikutuksia työnhakijan asemaan rekrytoinnissa

Tässä tutkielmassa tarkastellaan erityisesti tekoälyjärjestelmien vaikutuksia työnhakijan työnhakuprosessiin, sillä työnhakijat ovat tutkimuksen keskiössä. Näiden järjestelmien toiminnallisuus vaikuttaa suoraan tai epäsuorasti työnhakijoiden mahdollisuuksiin löytää työtä, joka vastaa heidän kykyjään tai tarpeitaan. Tekoäly voi tuoda mukanaan sekä hyötyjä, kuten prosessien tehokkuuden lisääntyminen, että riskejä, kuten päätöksenteon läpinäkyvyyden puute, mikä tekee tarkastelusta erityisen tärkeää.

Työnhakijoiden ideaalitalanne on löytää työtä organisaatioista, jotka vastaavat hyvin heidän kykyjään, palkkatasoaan, koulutustaustaansa ja sosiaalisia verkostoja (van Esch ja muut, 2019, s. 216). Hunkenschroer & Luetge (2022, s. 1 000) toteavat, että tekoäly ja algoritmiset päätöksentekojärjestelmät voivat joko helpottaa tai hankaloittaa työnhakijoiden tavoitteen saavuttamista rekrytoinnissa. Tämä johtuu siitä, kuinka hyvin tekoälyjärjestelmät todellisuudessa toimivat oikeudenmukaisuuden ja läpinäkyvyyden kannalta. Eettiset näkökulmat korostuvat erityisesti, kun tarkastellaan näiden järjestelmien vaikutusta oikeudenmukaisuuteen, läpinäkyvyyteen sekä inhimillisen valvonnan ja autonomian säilymiseen rekrytointiprosessissa (Hunkenschroer & Luetge, 2022, s. 993).

Tekoälyn automaattisen päätöksenteon ja optimoinnin vaikutukset korostuvat työnhakijan kannalta rekrytoinnin vaiheissa, kuten tavoittamisessa, seulonnassa, arvioinnissa ja koordinoinnissa (Hunkenschroer & Luetge, 2022, s. 992). Hunkenschroer & Luetge (2022, s. 992) toteavat, että työnhakijat ovat jatkuvan profiloinnin kohteena, ja heihin kohdistetaan viestintää algoritmisesti verkkoalustoilla ja sosiaalisessa mediassa tavoittavuuden lisäämiseksi. Tällä toiminnalla pyritään luomaan mahdollisimman laaja hakijapooli sekä tavoittamaan aktiivisten työnhakijoiden lisäksi myös passiivisia hakijoita nopeasti ja tehokkaasti (Johnson ja muut, 2021, s. 43).

Black & van Esch (2020, s. 220) tarkastelevat tavoittavuutta positiivisesta näkökulmasta, sillä tekoälyn katsotaan parantavan työnhakijoiden mahdollisuuksia löytää työpaikkoja, jotka vastaavat heidän taitojaan ja kiinnostuksiaan. Johnson ja muut (2021, s. 45) toteavat, että datan hyödyntäminen algoritmipohjaisessa viestinnässä mahdollistaa paremman ja tehokkaamman työnhakijoiden yhteensovittamisen organisaation työtarpeisiin tavoittamisvaiheessa. Digitaalisten lähtötietojen hyödyntäminen tekoälyllä nopeuttaa rekrytointiprosessia, mutta siihen voi liittyä rakenteellisia ja piileviä ongelmia, kuten datan puolueellisuus, algoritmien läpinäkymättömyys tai ihmiskontaktin puuttuminen (Hunkenschroer & Luetge, 2022, s. 999).

Data- ja algoritmikeskeinen rekrytointi tuo mukanaan monimutkaisia ongelmia, joihin on vaikea löytää yksiselitteisiä ratkaisuja datan käsittelyn monimuotoisuuden vuoksi. Hunkenschroer & Luetge (2022, s. 999) toteavat, että laajempi hakijapooli ja lisääntynyt kilpailu vaikuttavat työnhakijan mahdollisuuteen saada työpaikka. Datakeskeisyys ja sen seurauksena korostuva pisteytys painottuvat työnhakijoiden valinnassa. Lisäksi Albassam (2023, s. 10–11) käsittelee dataturvallisuuden huolia ja algoritmien vinoumia, jotka voivat luoda epätasa-arvoisia tilanteita työnhakijoiden kesken. Haresamudram ja muut (2023, s. 96) huomauttavat, että esimerkiksi Amazonin tarjoama tekoälypohjainen rekrytointijärjestelmä todettiin syrjivän naisia. Tekoälyn käyttö rekrytointin alkuvaiheessa voi nopeuttaa ja tehostaa päätöksentekoa, mutta kysymyksiä herää edelleen algoritmien läpinäkyvyydestä ja oikeudenmukaisuudesta, kuten puolueellisten lähtötietojen vaikutuksesta työnhakijoiden valintaan (Hunkenschroer & Luetge, 2022, s. 99).

Data- ja algoritmikeskeisyys korostuu seulontavaiheessa. Albaroud ja muut (2024, s. 391) toteavat, että tekoäly mahdollistaa ansioluetteloiden seulonnan automaatioprosessin, jonka avulla pyritään tunnistamaan parhaat työnhakijat. Ansioluetteloiden ja saatekirjeiden analysointi on tyypillinen tehtävä, jossa työnhakija arvioidaan lähettämien liitetiedostojen perusteella (Black & van Esch, 2020, s. 220). Lookadoo & Moore (2024, s. 14) huomauttavat, että algoritmien käyttö seulontavaiheessa perustuu muun muassa tekstin avainsanoihin ja lauserakenteisiin, mikä voi mahdollisesti johtaa työnhakijan sivuuttamiseen joko inhimillisistä tai teknisistä syistä.

Tietojen automaattinen kerääminen ansioluetteloista on haastavaa, sillä ne ovat usein jäsentämättömiä asiakirjoja, joissa on paljon vaihtelua tyylessä ja sisällössä (Palshikar ja muut, 2023, s. 12). Lookadoo & Moore (2024, s. 10) mukaan ongelma on lähtökohtaisesti monitasoinen. Yksi keskeinen ristiriita liittyy rekrytointiasiantuntijoiden neuvoihin ansioluetteloiden laadinnassa, sillä esimerkiksi suositellut fontit ja visuaaliset tyylit voivat heikentää koneluettavuutta ja siten alentaa hakijoiden pisteitä merkittävästi seulontavaiheessa. Hunkenschroer & Luetge (2022, s. 997) korostavat, että tämän lisäksi

rekrytointijärjestelmät ja organisaatiot jakavat työnhakijoille vain vähän tietoa siitä, miten heidän lähettämiään asiakirjoja, kuten ansioluetteloita ja saatekirjeitä, käsitellään. Tämä tarkoittaa, että hakijat eivät saa tietoa, jos heidän asiakirjojensa heikko koneluettavuus johtaa pistevähennyksiin ja pudottaa heidät rekrytointiprosessista ilman mahdollisuutta korjata tilannetta (Lookadoo & Moore, 2024, s. 10).

Arviointivaiheessa työnhakijoita tarkastellaan datalähtöisesti. Hunkenschroer & Luetge (2022, s. 992–993) toteavat, että datalähtöiset arviointimenetelmät mahdollistavat syvällisemmän ymmärryksen hakijoiden kyvyistä ja persoonallisuuksista. Tekoälypohjaiset työkalut, kuten videohaastattelut, testit ja psykologiset pelit, tarjoavat rekrytoijille arvokasta tietoa, joka laajentaa perinteisiä arviointikriteerejä (Black & van Esch, 2020, s. 220–221). Hunkenschroer & Luetge (2022, s. 996–997) toteavat, että tällainen läpinäkyvyys voi estää väärinkäsityksiä ja epäluottamusta, joita voi syntyä, kun hakijat eivät tiedä, miten heidän tietojaan käsitellään. Tämä ei ainoastaan paranna työnhakijoiden luottamusta rekrytointimenettelyihin, vaan myös antaa heille mahdollisuuden vaikuttaa omaan arviointiinsa ja profiiliinsa (Van Iddekinge ja muut, 2023, s. 667).

Koordinointi rekrytointiprosessin eri vaiheissa on keskeinen tekijä työnhakijakokemuksen ja -aseman kannalta. Hunkenschroer & Luetge (2022, s. 997) toteavat, että tekoäly voi tehostaa rekrytoinnin koordinointia ja parantaa viestintää rekrytoijien ja työnhakijoiden välillä, vähentäen viivästyksiä ja parantaen hakijakokemusta. Kuitenkin automaattinen koordinointi voi luoda epätasa-arvoisia tilanteita, sillä algoritmit saattavat suosia hakijoita, jotka vastaavat tarkasti aiempia menestyneitä profiileja, jättäen potentiaalisesti hyviä ehdokkaita huomiotta. Tehokas koordinointi voi myös vaikuttaa työnhakijoiden itsetuntoon ja motivaatioon, sillä hakijat, jotka saavat palautetta ja tukea prosessin aikana, kokevat hakukokemuksensa todennäköisesti myönteisempänä (Van Iddekinge ja muut, 2023, s. 667). Siksi organisaatioiden on tärkeää panostaa koordinaation läpinäkyvyyteen ja varmistaa, että työnhakijat saavat tarvittavaa tietoa ja palautetta koko rekrytointiprosessin ajan. Alla

olevassa taulukossa 1 on tiivistetysti esitelty tekoälyn vaikutuksia työnhakijan asemaan hyötyjen ja riskien kannalta.

| Hyödyt | Riskit |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> • Prosessin nopeutuminen • Laajempi tavoitavuus • Syvällisempi arviointi hakijan taidoista ja persoonallisuuksista • Viestinnän automaatio ja työnhakijan kokema helppous tavoittamisessa. | <ul style="list-style-type: none"> • Läpinäkyvyyden puute, algoritmien puolueellisuus • Työnhakijan vaikutusmahdollisuuksien väheneminen • Inhimillisen kontaktin puute • Vähemmän mahdollisuuksia palautteeseen • Ansioluetteloiden koneluetavuusongelmat |

Taulukko 1. Vaikutuksista työntekijän asemaan hyötyjen ja riskien kannalta (Black & van Esch, 2020; Hunkenschroer & Luetge, 2022; Lookadoo & Moore, 2024).

3 Selittävä tekoäly

Selittävä tekoäly, eli XAI (eng. Explainable Artificial Intelligence), on tekoälyn ala, joka keskittyy tarjoamaan ymmärrettäviä selityksiä älykkäiden järjestelmien tekemille ennusteille, suosituksille ja päätöksille (Ridley, 2022, s. 2). Chromik & Butz (2021, s. 2) esittävät, että selitettävän tekoälyn lähestymistapa voidaan yksinkertaisesti jakaa kahteen päävaiheeseen, jotka ovat kone- ja syväoppimismallien kehittäminen ja selityksiä tuottavan käyttöliittymän suunnittelu. Dwivedi ja muut (2023, s 8–9) toteava, että kone- ja syväoppimisen mallien analysointi sekä kehitetyt selittävät menetelmät voivat parantaa ymmärrystä tekoälyn päätöksentekoprosesseista ja tuloksista algoritmisella tasolla. Kone- ja syväoppimismalleja on monia erilaisia tiedonkäsittelyn tarkoituksia varten ja niiden kehittämisellä pyritään tuottamaan selityksiä tekoälyjärjestelmien käyttökontekstien ja tarpeiden mukaisesti. Selittävän tekoälyn lähestymistapa on laajentumassa teknokeskeisestä algoritmien läpinäkyvyydestä kohti monitieteellisempää konseptia, jossa huomioidaan eri sidosryhmien odotuksia ja tarpeita tekoälyn läpinäkyvyyden suhteen vuorovaikutuksellisella ja sosiaalisella tasolla (Haresamudram ja muut, 2023, s. 98–99).

Ihmisten yksilölliset käsitykset, tekninen monimutkaisuus ja tekoälyjärjestelmien moninaiset käyttökontekstit luovat ongelmia selittävän tekoälyn ratkaisujen yleisluonteisten määritelmien ja viitekehyksien kehittämiseksi. Vuorovaikutukselliset, sosiaaliset ja poliittiset ulottuvuudet lisäävät tähän vielä oman kerroksensa haasteita (Haresamudram ja muut, 2023, s. 99). Ridley (2022) mukaan Yhdysvaltain puolustusministeriön alaisen DARPA-organisaation (eng. The US Defense Advanced Research Projects Agency) kuvaus selitettävästä tekoälystä tarjoaa kattavan ja käyttäjakeskeisen lähestymistavan. Selitettävän tekoälyn tarkoituksena on, että tekoälyjärjestelmillä on kyky selittää perustelunsa sekä kuvata vahvuutensa ja heikkoutensa, jotta käyttäjien ymmärrys, luottamus ja hallinta optimoituvat yhteistyössä tekoälyratkaisujen käyttäjien kanssa. Yleistavoitteena on edistää tekoälyjärjestelmien läpinäkyvyyttä ja vastuullista käyttöä yksilöistä organisaatioihin ja sääntelyelimiin, jotta

voidaan luoda tehokkaita ja eettisiä tekoälyratkaisuja, jotka täyttävät sekä tekniset että yhteiskunnalliset vaatimukset (Gunning & Aha, 2019, s. 56–57).

3.1 Läpinäkyvyyden merkitys

Läpinäkyvyys on moninainen käsite, jota on vaikea määrittellä yksiselitteisesti (Haresamudram ja muut, 2023, s. 94). Felzmann ja muut (2020, s. 3 335) toteavat, että läpinäkyvyyden perusta on informaation välittämisessä, johon sisältyvät selitettävyys, avoimuus, saavutettavuus ja näkyvyys. Pelkkä informaation välittäminen on kuitenkin riittämätöntä, sillä läpinäkyvydessä huomioidaan myös ihmisten sosiaaliset tekijät, arvot ja informaation merkitys, jotta se voidaan ymmärtää paremmin (Felzmann ja muut, 2020, s. 3 336).

Lähtökohtaisesti läpinäkyvydellä pyritään luomaan positiivista arvoa, mutta ongelmana on määrittää, mitä se eri sidosryhmille tarkoittaa ja kuinka hyödyllistä se on (Felzmann ja muut, 2020, s. 3 336). Felzmann ja muut (2020, s. 3 336) toteavat, että käsitys läpinäkyvyydestä hyveenä on normatiivinen käsite, joka asettaa standardit julkisten toimijoiden käyttäytymisen arvioimiselle. Läpinäkyvyys hyveenä nähdään luontaisesti positiivisena arvona, joka koostuu toimijoiden, järjestelmien ja organisaatioiden johdonmukaisesta avoimuudesta toiminnassaan, käyttäytymisessään, aikomuksissaan ja harkinnoissaan (Schmidt & Wood, 2019, s. 728–729). Kuitenkin Felzmannin ja muiden (2020, s. 3 336) mukaan pelkkä läpinäkyvyyden käsittäminen hyveenä ei riitä, sillä se ei määrittele yleisöä tai kohdetta, jolle toimija on läpinäkyvä.

Suhteellisessa läpinäkyvydessä läpinäkyvyys käsitetään suhteena toimijan ja informaation vastaanottajan välillä (Rego ja muut, 2022, s. 696). Bovensin ja muiden (2014) mukaan on tärkeää huomata, että läpinäkyvyys ei ole pelkästään yksilöllinen ominaisuus, vaan myös vuorovaikutussuhde, jossa toimijoiden avoimuuden vastaanotto ja ymmärtäminen ovat olennaisia. Suhteellisessa näkökulmassa keskeistä on, että läpinäkyvyystoimenpiteet suunnitellaan ja arvioidaan niiden vaikutusten perusteella eri sidosryhmiin (Felzmann ja muut, 2020, s. 3 336).

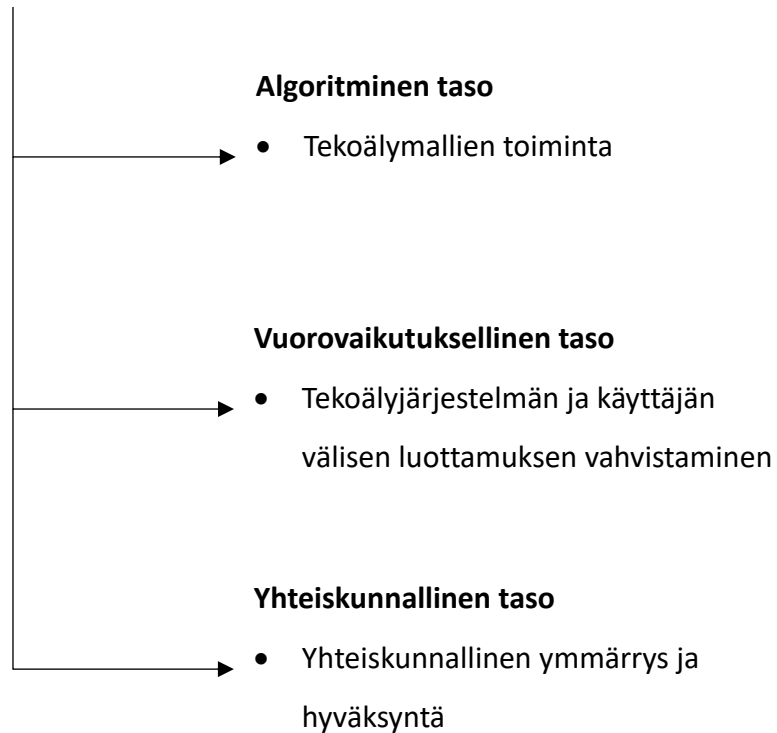
Suhteellisessa läpinäkyvydessä on huomioitava myös suhteiden institutionaalinen konteksti (Felzmann ja muut, 2020, s. 3 336). Felzmann ja muiden (2020, s. 3 337) mukaan institutionaalisen kontekstin erityispiirteet, kuten oikeudelliset, sääntelylliset ja organisatoriset toimenpiteet, on huomioitava, jotta läpinäkyvyyden käytännön vaikutukset ja tehokas toteutus voidaan realistisesti ymmärtää. Euroopan tietosuojalain keskeinen periaate on läpinäkyvyys, mikä sisältää standardit sille, miten henkilötietoja käsitellään tietyissä konteksteissa, kuten esimerkiksi rekrytinnissa, joka on henkilötietojen osalta sensitiivistä aluetta (Schmidt & Wood, 2019, s. 729).

3.1.1 Läpinäkyvyyden tasot tekoälyjärjestelmissä

Tekoälyjärjestelmiin liittyvät sidosryhmät ja niiden suhteet järjestelmiin voidaan luokitella monipuolisesti intressien ja käyttötarkoitusten perusteella (Langer ja muut, 2021, s. 3). Haresamudramin ja muiden (2023, s. 95) mukaan selittävässä tekoälyssä läpinäkyvyyden tasot voidaan luokitella algoritmiseksi, vuorovaikutukselliseksi ja yhteiskunnalliseksi. Algoritminen läpinäkyvyys viittaa kone- ja syväoppimisen mallien kehittämiseen, vuorovaikutuksellinen läpinäkyvyys tekoälyjärjestelmän ja sen käyttäjien luottamuksen vahvistamiseen ja sosiaalinen läpinäkyvyys kattaa laajemman yhteiskunnallisen ymmärryksen ja hyväksynnän (Haresamudram ja muut, 2023, s. 96–98). Alla olevassa kuviossa 5 on havainnollistettuna kolme läpinäkyvyyden tasoa tekoälyjärjestelmissä.

Läpinäkyvyyden tasot

tekoälyjärjestelmissä



Kuvio 5. Läpinäkyvyyden tasot tekoälyjärjestelmissä (Haresamudram ja muut, 2023, s. 95).

Algoritmisessa läpinäkyvydessä käsitellään tekoälymalleja, kuten kone- ja syväoppimismalleja (Haresamudram ja muut, 2023, s. 96). Campesato (2020, s. 18–19) kuvaa kone- ja syväoppimismallit tekoälyn alaluokkina, jotka mahdollistavat tehtävien ratkaisemisen, jotka olisivat perinteisillä ohjelmointikielillä joko mahdottomia tai liian monimutkaisia. Koneoppimismallit oppivat itsenäisesti tunnistamaan kaavoja ilman tarkasti määriteltyä ohjelmointia, tekemään päätöksiä datan perusteella ja mukautumaan uusiin tilanteisiin (Alexander Jung, 2022, s. 12). Janiesch (2021, s. 687) mukaan syväoppimismallit puolestaan hyödyntävät monikerroksisia hermoverkkoja

monimutkaisempien kaavojen ja kuvioden tunnistamiseksi, mikä mahdollistaa korkeamman tarkkuuden ja tehokkuuden monimutkaisissa tehtävissä.

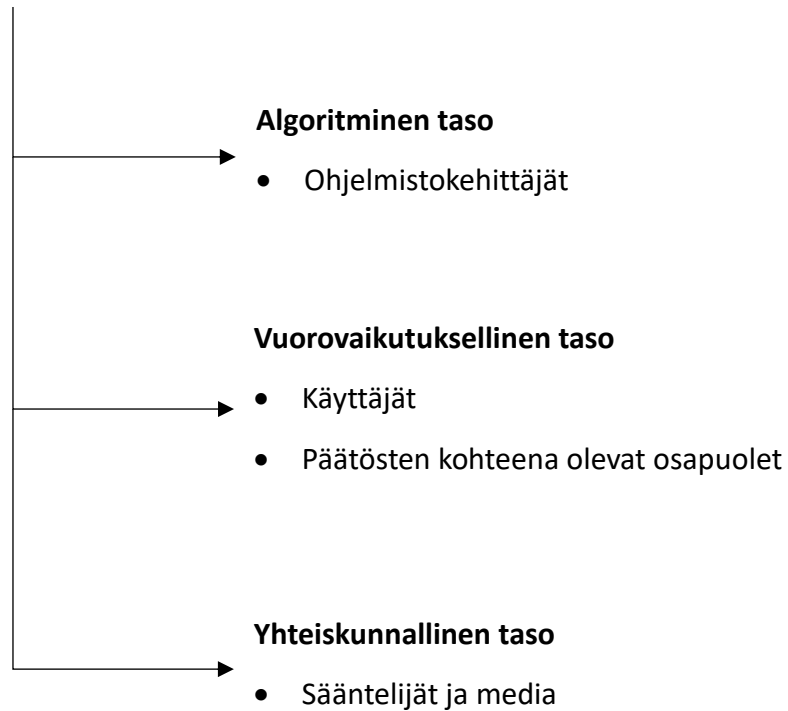
Vuorovaikutuksellisessa läpinäkyvydessä keskitytään siihen, miten tieto vaihtuu tekoälyjärjestelmän ja käyttäjien välillä vuorovaikutuksen kautta (Haresamudram ja muut, 2023, s. 97). Miller (2019, s. 32–33) toteaa, että kyse on tarkemmin ihmisen ja tekoälyagentin tai -ohjelmiston välisestä vuorovaikutuksesta, johon sisältyy yhteistyötä ja kommunikointia. Wiberg & Stolterman Bergqvist (2023, s. 2 289) mukaan tekoälyjärjestelmien kehitys suuntautuu vahvasti päätöksentekoprosessien automatisoitumiseen. Tämä johtaa vähäisempään vuorovaikutukseen käyttäjien kanssa, sillä järjestelmät kykenevät yhä itsenäisemmin suorittamaan monimutkaisia tehtäviä ilman ihmisen väliintuloa. Toisaalta tekoälyjärjestelmien automaatiot sekä kyky oppia ja mukautua mahdollistavat käyttäjille uusia vuorovaikutusparadigmoja, kuten selitysten tuottamisen päätöksentekoprosessista ja tuloksista algoritmisiin toimintoihin (Haresamudram ja muut, 2023, s. 97).

Yhteiskunnallisessa läpinäkyvydessä keskitytään tekoälypalveluntarjoajiin, lainsäädännöllisiin kehyksiin ja tekoälyn vaikutuksiin sosiaalisessa kontekstissa (Haresamudram ja muut, 2023, s. 95). Ryanin & Stahlin (2021) mukaan tekoälyn läpinäkyvyys on keskeinen tekijä sen hyväksyttävyyden kannalta yhteiskunnassa. Läpinäkyvyysvaatimusten nousu lainsäädännöllisissä kehyksissä ja sosiaalisissa konteksteissa perustuu teknologiayritysten tarjoamien tekoälypalveluiden kriittiseen tarkasteluun (Hollanek, 2023, s. 2 079). Hollanek (2023, s. 2 077) toteaa, että kriittinen tarkastelu on ohjannut teknologiayrityksiä muuttamaan toimintatapojaan ja panostamaan algoritmisen läpinäkyvyyden kehittämiseen. Joissakin palveluissa on havaittu ristiriitoja oletetun algoritmisen toiminnan ja niiden todellisen vaikutuksen välillä käyttäjien kannalta. Selittävän tekoälyn periaatteiden ja käytäntöjen on osoitettu edistävän tekoälyn läpinäkyvyyttä ja ymmärrettävyyttä monissa sovelluksissa yhteisten kriteerien mukaisesti (Barredo Arrieta ja muut, 2020, s. 107–108).

3.1.2 Sidosryhmien intressit läpinäkyvyyden tasoilla

Ihmisyryhmillä on erilaisia intressejä tekoälyjärjestelmien selitettävyyden suhteen, kuten järjestelmän kehittäminen, päätöksiin liittyvän tiedon tuottaminen, työtehtävien suorittaminen tai sääntelykehityksen luominen (Langer ja muut, 2021, s. 3). Barredo Arrieta ja muut (2020) esittelevät sidosryhmien luokittelun, johon kuuluvat käyttäjät, ohjelmistokehittäjät, tekoälyjärjestelmien päätösten kohteena olevat osapuolet, järjestelmän käyttöönottajat ja sääntelijät. Haresamudram ja muut (2023, s. 96) lisäävät tähän luokitukseen yhteiskunnan muita toimijoita ja suunnittelijoita, kuten mediatoimijat ja vuorovaikutussuunnittelijat. Tekoälyperusteisissa henkilöstöhallinta- tai rekrytointijärjestelmissä sidosryhmät voidaan luokitella pääsääntöisesti ohjelmistokehittäjiin, käyttäjiin, päätösten kohteena oleviin osapuoliin, käyttöönottoihin, sääntelijöihin ja suunnittelijoihin (Langer ja muut, 2021, s. 4). Alla oleva kuvio 6 esittää yleisen sidosryhmien luokituksen läpinäkyvyyden tasoilla.

**Sidosryhmien intressit
läpinäkyvyyksien tasolla**



Kuvio 6. Sidosryhmien luokitus läpinäkyvyyden tasoilla (Barredo Arrieta ja muut, 2020, s. 84).

Ohjelmistokehittäjät voidaan luokitella tarkemmin kolmannen osapuolen kehittäjiin ja organisaatioiden sisäisiin kehittäjiin (Langer & König, 2023, s. 6). Haresamudram ja muiden (2023, s. 96) mukaan ohjelmistokehittäjien ensisijainen intressi on algoritminen läpinäkyvyys. Algoritminen läpinäkyvyys tarkoittaa sitä, kuinka avoimesti ja ymmärrettävästi algoritmin toiminta ja päätöksentekoprosessit esitetään (Dwivedi ja muut, 2023, s. 8). Haresamudram ja muut (2023, s. 96) arvioivat, että tekoälyjärjestelmien kannalta algoritmisen läpinäkyvyyden täydellinen saavuttaminen voi olla kehittäjille ongelmallinen, sillä suuret tietomäärät ja monimutkaiset mallit voivat johtaa ennakoimattomiin ja tuntemattomiin lopputuloksiin. He toteavat, että

tekoälyjärjestelmien päätöksentekoprosessit voivat olla kokonaisuudessaan ihmismielelle hallitsemattomia. Algoritminen läpinäkyvyys on kuitenkin erityisen tärkeää ohjelmistokehittäjille, sillä se mahdollistaa kriittisen tarkastelun ja varmistaa, että ne toimivat suunnitellulla tavalla ja tuottavat luotettavia tuloksia (Dwivedi ja muut, 2023, s. 8).

Langer & König (2023, s. 4) esittävät, että tekoälypohjaisten henkilöstöhallinta- ja rekrytointijärjestelmien käyttäjät voidaan luokitella tarkemmin henkilöstöhallinnon johtajiin ja työntekijöihin. Tekoälyn ja käyttäjän välinen yhteistyö vaikuttaa erityisesti vuorovaikutuksellisen ja sosiaalisen läpinäkyvyyden tasolla (Haresamudram ja muut, 2023, s. 98). Langer & König (2023, s. 4) kuvailevat, että henkilöstöhallinta- ja rekrytointijärjestelmien vuorovaikutuksellisen läpinäkyvyyden tavoitteena on käyttäjien päätöksenteon tehokkuuden ja laadun parantaminen henkilöstöhallintaan liittyvissä tehtävissä. Barredo Arrieta ja muut (2020, s. 86) korostavat, että selittävän tekoälyn keinoin pyritään esittämään käyttäjille päätöksentekoprosessin syy-seuraussuhteita ja lopputuloksia informatiivisesti, jotta käyttäjät voivat arvioida päätösten oikeudenmukaisuutta ja luotettavuutta paremmin. Tekoälyjärjestelmien läpinäkyvyys ja todettu toimivuus sitouttavat käyttäjiä, mikä lisää järjestelmien sosiaalista hyväksyntää organisaatiotasolla (Langer & König, 2023, s. 4).

Päätöksien kohteena olevat ihmiset eivät yleensä voi valita, että vaikuttaako algoritmipohjainen päätöksenteko heihin (Langer & König, 2023, s. 5). Langer & König (2023, s. 5) toteavat, että henkilöstöhallinta- ja rekrytointijärjestelmien päätösten kohteena olevat osapuolet voidaan luokitella työnhakijoihin ja organisaation muihin työntekijöihin, joihin vaikuttavat valintapäätökset, suoritusarvioinnit ja aikataulutukset. Myös käyttäjät, eli henkilöstöhallinnon johtajat ja työntekijät, voidaan lukea päätöksen kohteena oleviin osapuoliin, sillä heihin vaikuttavat tekoälyjärjestelmän tuottamat päätökset, joita he hyödyntävät työsuorituksissaan. Päätösten kohteena olevat osapuolet sijoittuvat vuorovaikutuksellisen ja sosiaalisen läpinäkyvyyden tasolle, vaikka heidän vuorovaikuttamismahdollisuutensa tekoälyjärjestelmien tuottamien päätösten,

kuten rekryointiarviointien, suhteen ovat usein rajoitettuja (Haresamudram ja muut, 2023, s. 98).

Tekoälyjärjestelmän käyttöönottajat voidaan luokitella kahteen alaryhmään, jotka ovat algoritmipohjaisia henkilöstöhallinta- tai rekryointijärjestelmiä tarjoavat organisaatiot ja organisaatioiden päätöksentekijät, joiden vastuulla on käyttöönotto sekä soveltaminen prosesseihin (Langer & König, 2023, s. 5). Felzmannin ja muiden (2020, s. 7–10) mukaan käyttöönottajilla on merkittävä rooli sosiaalisen läpinäkyvyyden varmistamisessa organisaatioiden strategisella tasolla. Organisaatiot käytännössä määrittelevät, mikä on läpinäkyvää, ja nämä päätökset vaikuttavat myös muihin sidosryhmiin. Tämä vaikuttaa erityisesti ohjelmistokehittäjiin, sillä organisaatioiden päätöksentekijät määrittävät, miten järjestelmien läpinäkyvyyttä on tarkoitus toteuttaa algoritmisella tasolla ja mitä siitä näytetään muille sidosryhmille. Läpinäkyvyyteen liittyvät valinnat ja päätökset vaikuttavat lopulta monilla tasoilla, kuten yrityksen kilpailuetuun, päätösten kohteena olevien ihmisten mielikuviin päätöksentekoprosessista sekä sääntelijöiden arvioihin oikeudenmukaisuudesta (Barredo Arrieta ja muut, 2020, s. 84).

Goodmanin & Flaxmanin (2017, s. 55–56) mukaan sääntelijöiden valvontarooli tekoälyn suhteen on ongelmallinen. Tavoitteena on saavuttaa korkea tehokkuus ja hyötyjä kaikille rekryointiprosessin sidosryhmille, vaikka joissakin tekoälyperusteisissa rekryointijärjestelmissä on karsittu läpinäkyvyydestä. Sääntelyyn liittyviä ongelmallisia kohtia ovat organisaatioiden immateriaalioikeuksien suoja, järjestelmien auditointi ja vastuun tarkka määrittäminen (Langer & König, 2023, s. 6). Goodman & Flaxman (2017, s. 56) toteavat, että oikeus syrjimättömyyteen ja selitykseen asettaa tekoälyn kehittäjille, sääntelijöille ja tieteenaloille positiivisen ongelman läpinäkyvyyden parantamiseksi tulevaisuudessa. Erityisesti Euroopassa toimivien organisaatioiden on panostettava läpinäkyvyyden kehittämiseen, sillä Euroopan unionin tietosuojasetus ja tekoälysäädös kehittyvät tarkempien standardien ja menetelmien määrittäjäksi (Langer & König, 2023, s. 6).

Suunnittelijat voidaan luokitella tarkemmin ihmisen ja teknologian välisen vuorovaikutuksen asiantuntijoiksi, jotka keskittyvät käyttöliittymien vuorovaikutussuunnitteluun (Langer & König, 2023, s. 6). Wiberg & Stolterman Bergqvist (2023, s. 2 288–2 289) toteavat, että tekoälyn mahdollistama automaatio vähentää manuaalista vuorovaikutusta järjestelmissä. Tämä ohjaa suunnittelijat tilanteisiin, joissa vuorovaikutuksen ja automaation välisen suhteen syvällinen ymmärtäminen on entistä tärkeämpää, sillä automaatio on käyttäjille usein näkymätön prosessi taustalla. On olennaista tunnistaa, miten käyttäjän vuorovaikutus vaikuttaa tekoälyjärjestelmän automaattisiin toimintoihin sekä mitä seurauksia tällä yhteistyöllä voi olla. Lisäksi on tärkeää asettaa vaatimuksia, jotka varmistavat järjestelmän luotettavan ja tehokkaan toiminnan. Suunnittelijoiden vaikutusmahdollisuudet vuorovaikutukseen voivat olla rajalliset, kun tekoäly pystyy sopeutumaan käyttäjiinsä automaattisesti. Tämä voi tarkoittaa, että järjestelmä mukauttaa vuorovaikutusominaisuuksiaan ilman suunnittelijan panosta (Haresamudram ja muut, 2023, s. 97–98).

3.1.3 Selittävän tiedon ymmärtäminen

Selittävä tieto auttaa ihmisiä ymmärtämään, ennustamaan ja hallitsemaan tilanteita kaventamalla tapahtumien mahdollisia syitä, vähentämällä epävarmuutta, korjaamalla väärinkäsityksiä ja mahdollistamalla syy-seuraussuhteiden hahmottamisen (Miller, 2019, s. 4–6). Langer ja muut (2021) toteavat, että Tekoälyn kontekstissa selittävä tieto tukee sidosryhmiä järjestelmän toiminnan ymmärtämisessä, epävarmuuden vähentämisessä ja odottamattomien tulosten erottamisessa järjestelmävirheistä. Selityksen käsitettä on tutkittu laajasti sekä filosofisesta että psykologisesta näkökulmasta, joissa korostuu kausaalisuuden merkitys (Miller, 2019, s. 6). Tämä syy-seuraussuhteiden ymmärtäminen muodostaa perustan selityksen kyvylle vähentää epävarmuutta ja tuoda selkeyttä monimutkaisiin tilanteisiin.

Kausaalinen päättely on keskeinen prosessi monille korkeatasoisille kognitiivisille toimintoille, jotka ovat erityisen tärkeitä sosioteknisissä järjestelmissä (Hoffman & Klein,

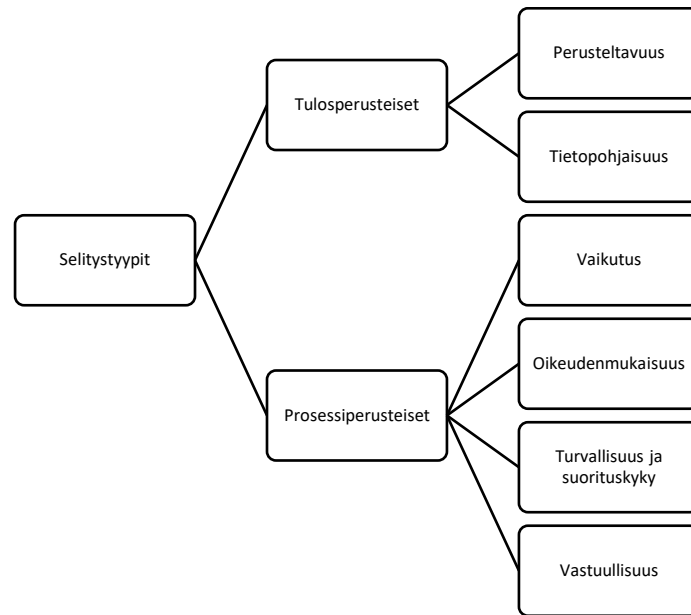
2017, s. 68). Hoffman & Klein (2017) toteavat, että tämän avulla pyritään ymmärtämään syy-seuraussuhteita sekä selittämään, miten tietyt tapahtumat, toimenpiteet tai ilmiöt johtavat tiettyihin lopputuloksiin. He tutkivat selityksen luonnetta ja tapoja, joilla asioita voidaan selittää ja esittää älykkäissä järjestelmissä, jotta niiden toimintaansa on ymmärrettävämpää. Kausaalinen päättely pohjautuu teoreettisiin perusteisiin ja jakautuu kolmeen päättelymuotoon, jotka ovat abduktiivinen, keskittyen todennäköisimpiin selityksiin, retrospektiivinen ja kontrafaktuaalinen, joissa analysoidaan menneitä tapahtumia ja vaihtoehtoisia skenaarioita, sekä prospektiivinen, joka ennakoii mahdollisia tulevia lopputuloksia (Hoffman & Klein, 2017, s. 71–71).

Abduktiivinen päättely on päättelyn muoto, jossa pyritään selittämään havainnot muodostamalla paras mahdollinen hypoteesi, joka selittää kyseiset havainnot (Hoffman & Klein, 2017, s. 69–70). Dreamson & Khine (2022, s. 405–406) korostavat, että tämä on prosessinomaista ongelmanratkaisua, jossa syitä etsitään sääntöjen ja vaikutusten perusteella. Heidän mukaansa abduktiivinen päättely sopii erityisesti suunnittelutieteellisiin ongelmanratkaisuihin, sillä se on luonteeltaan dialoginen, kriittinen, argumentoiva ja avoin uusille todisteille. Koska abduktiivista päättelyä voidaan käyttää sekä menneisyyden selittämiseen että tulevaisuuden ennakkointiin, se yhdistää retrospektiivisen ja prospektiivisen lähestymistavan (Hoffman & Klein, 2017, s. 70).

Retrospektiivinen ja kontrafaktuaalinen päättely keskittyy menneisyyden tapahtumien selittämiseen ja siihen, mitä olisi voinut tapahtua, jos tietyt olosuhteet olisivat erilaisia (Hoffman & Klein, 2017, s. 70). Hoffmanin & Kleinin (2017, s. 71–72) esittävät, että prospektiivinen kausaalinen päättely tarkastelee asioita, jotka voivat tapahtua tai jäädä tapahtumatta nykyhetkessä, ja jotka saattavat johtaa tai olla johtamatta tulevaisuuden seurauksiin. Abduktion malli sekä retrospektiivinen ja mentaalinen, tulevaisuuteen suuntautuva projektiivinen ajattelu muodostavat kausaalisen päättelyn teorian ja luonnollisen selitysmallin perustan (Hoffman & Klein, 2017, s. 78).

Tekoälyjärjestelmän kuusi keskeistä selitystyyppiä käyttäjille ovat perusteleva, vastuullinen, tietopohjainen, oikeudenmukainen, turvallisuus- ja suorituskky ja vaikutus (Leslie & Briggs, 2021, s. 11). Leslie & Briggs (2021, s. 12) luokittelevat selitystyyppit tarkemmin prosessi- ja tulosperusteisiin jakaumiin. Prosessiperusteisia selitystyyppijä ovat vastuullinen, oikeudenmukainen, turvallisuus- ja suorituskky ja vaikutus, kun taas tulosperusteisia ovat perusteleva ja tietopohjainen (Leslie & Briggs, 2021, s. 12).

Leslie & Briggs (2021, s. 11) korostavat, että perustelevassa selitystyyppissä kuvataan päätökseen johtaneet syyt ymmärrettävällä ja ei-teknisellä tavalla. Vastuullisessa selitystyyppissä esitetään, keitä on mukana tekoälyjärjestelmän kehittämisessä ja hallinnassa, sekä kehen voidaan olla yhteydessä ihmistarkastusta varten (Felzmann ja muut, 2020, s. 9). Barredo Arrieta ja muut (2020, s. 85) määrittelevät, että tietopohjainen selitystyyppi sisältää kuvauksen käytetyistä tiedoista ja niiden soveltamisesta päätöksenteossa. Oikeudenmukaisessa selitystyyppissä keskitytään päätösten puolueettomuuteen yksilötasolla (Langer ja muut, 2021, s. 11–12). Miller (2019, s. 24–25) Miller (2019, s. 24–25) painottaa, että turvallisuus- ja suorituskky selitystyyppi on toimenpidekeskeinen ja esittää, miten suunnittelulla ja toteutuksella pyritään maksimoimaan tarkkuus, luotettavuus ja turvallisuus. Vaikutusselitystyyppi puolestaan käsittelee toimenpiteitä, joilla arvioidaan järjestelmän suunnittelua ja toteutusta sekä seurataan sen käytön ja päätösten vaikutuksia yksilöihin ja yhteiskuntaan (Leslie & Briggs, 2021, s. 11). Alla oleva kuvio 7 havainnollistaa selitystyyppit.



Kuvio 7. Selitystyyppit (Leslie & Briggs, 2021, s. 12).

3.2 Selitettävyysetietoinen suunnittelu

Käyttäjälähtöisessä suunnitteluprosessissa, eli HCD-suunnittelussa (eng. Human-Centered Design), keskitytään käyttäjien tarpeisiin ja vaatimuksiin (Schoonderwoerd ja muut, 2021, s. 2). Schoonderwoerd ja muut (2021, s. 2) määrittelevät, että suunnitteluprosessi painottaa erityisesti sidosryhmien osallistamista, iteratiivista parantamista ja ihmisen suorituskvyn arviointia järjestelmissä. HCD-suunnitteluprosessi sisältää järjestelmän käsitteellisen ja fyysisen suunnittelun (Benyon, 2019, s. 51–52).

Käsitteellinen suunnittelu kattaa palvelun tai järjestelmän yleisen tarkoituksen ja vision, kun taas fyysinen suunnittelu konkretisoi abstraktit esitykset teknologisten vaatimusten ja ihmisten tehtävien kautta (Benyon, 2019, s. 51). Shneidermanin ja muiden (2018, s. 106) määrittelevät käsitteellisen suunnittelun mentaaliksi malleiksi, jotka kuvaavat, miten toiminnallisuudet ja ominaisuudet kommunikoivat käyttäjien kanssa. Fyysinen suunnittelu tarkoittaa teknologisia vaatimuksia sekä määrittää tehtävien ja roolien jakamisen ihmisten ja laitteiden välillä (Benyon, 2019, s. 51). Benyon (2019, s. 53) korostaa, että fyysisen suunnittelun tavoitteena on muuntaa abstraktit käsitteet

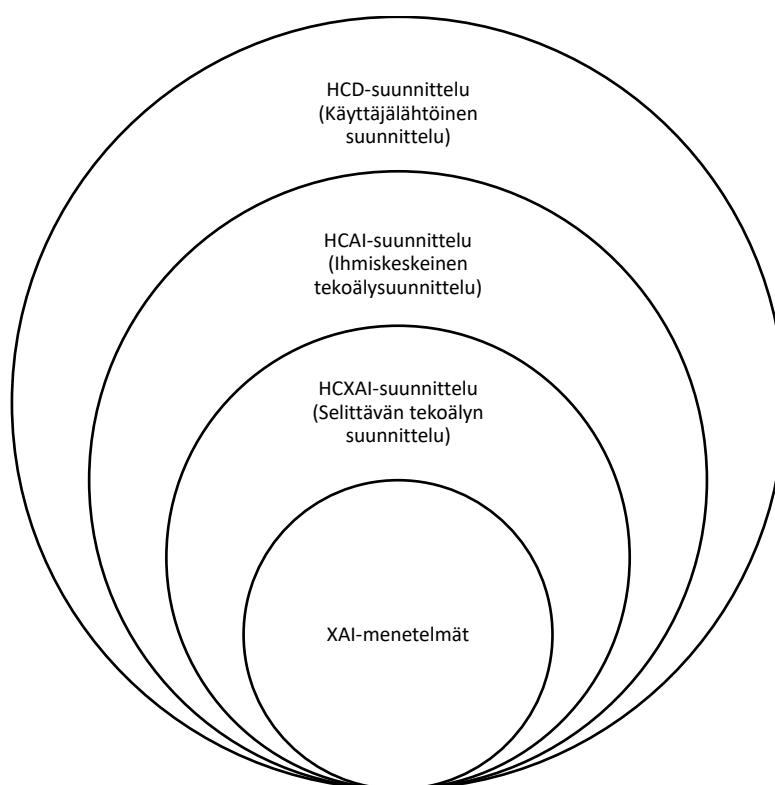
käytännön toteutettaviksi ratkaisuiksi, kuten prototyypeiksi ja käyttöliittymäsuunnitelmiksi.

Ihmiskeskeisessä tekoälysuunnittelussa, eli HCAI-suunnittelussa (eng. Human-Centered Artificial Intelligence), keskitytään käyttäjälähtöisiin suunnitteluprosesseihin, mutta huomioidaan myös tekoälyjärjestelmien automaattiset ominaisuudet (Shneiderman, 2022, s. 9). Shneidermanin (2022, s. 9) mukaan suunnittelussa huomioidaan erityisesti tekoälyjärjestelmien ominaisuuksien kykyä laajentaa ja parantaa ihmisen suorituskykyä, luovuutta, vastuullisuutta ja sosiaalisia yhteyksiä. Vaikka tekoäly yleistyy monimutkaisissa ja arkaluonteisissa socioteknisissä järjestelmissä, niiden tietojenkäsittely, päätöksentekoprosessit ja automaattiset ominaisuudet sekä toiminnallisuudet ovat usein käyttäjille läpinäkymättömiä (Bingley ja muut, 2023, s. 6).

HCAI-tutkijat ja suunnitteluohjeistukset korostavat, että tekoälyn tulisi olla tulevaisuudessa selitettävissä, ymmärrettävissä ja tulkittavissa kaikille käyttäjille ja niille, joihin tekoäly vaikuttaa (Bingley ja muut, 2023, s. 6). Ehsan & Riedl (2020) osoittavat, että ihmiskeskeiseen tekoälysuunnitteluun voidaan sisällyttää selittävän tekoälyn suunnittelumenetelmiä, jotka tukevat käyttäjiä tekoälyjärjestelmien ymmärtämisessä. Tätä lähestymistapaa kutsutaan ihmiskeskeiseksi selittäväksi tekoälyksi, eli HCXAI (eng. Human-Centered Explainable Artificial Intelligence). HCXAI-suunnitteluperiaatteessa keskitytään sociotekniseen lähestymistapaan, jossa pyritään kokonaisvaltaiseen ymmärrykseen ihmisestä, kuten ihmisten arvoista, ihmissuhteiden dynaamisuudesta ja tekoälyjärjestelmien sosiaalisesta kontekstista (Ehsan & Riedl, 2020, s. 12–13).

Selittävän tekoälyn suunnitteluperiaatteet, eli HCXAI-suunnitteluperiaatteet, seuraavat sovellussuunnittelun menetelmiä ja prosesseja, mutta niissä korostetaan erityisesti selitysmenetelmien kehittämistä ja käyttäjäkokemusta tekoälysovelluksen käyttötarkoituksen mukaisesti (Langer ja muut, 2021, s. 15–16). Clement ja muut (2023) esittävät kokonaisvaltaisen XAI-suunnittelumallin tekoälysovelluksille, joka on sisältää sovellussuunnittelun tyypilliset vaiheet, kuten vaatimusten analysointi, suunnittelu,

toteutus, arviointi ja kehitys. Jokaisessa vaiheessa keskeinen näkökulma keskittyy kuitenkin selitystietoiseen suunnitteluun. XAI-suunnitteluperiaatteissa korostetaan erityisesti XAI-menetelmät, sillä ne ovat tekoälymalliin liitettäviä menetelmiä, joilla tekoälymallien tuottamat tulosteet ja ennusteet muutetaan informaatioksi, josta konstruoidaan selityksiä käyttäjärühmille (Schoonderwoerd ja muut, 2021, s. 3–4). Alla oleva kuvio 8 havainnollistaa hiarkisen ja käyttäjäkeskeisen eri tasoja tekoälysuunnittelussa.



Kuvio 8. Suunnittelulähestymistapojen hierarkia (Benyon, 2019; Bingley ja muut, 2023).

3.2.1 Selitettävyystavotteet

Selittävän tekoälyn selitettävyystavotteet voidaan jakaa neljään näkökulmaan, jotka ovat mitä, miksi, mitä varten ja miten (Barredo Arrieta ja muut, 2020, s. 84). Gunningin & Ahan (2019) huomauttavat, että vaikka filosofiset teorit selitysten monimutkaisuudesta ja kriittisestä tarkastelusta ovat kiinnostavia, selittävän tekoälyn kannalta olennaisinta on kehittää algoritmisia tekniikoita, jotka tukevat tekoälyn

ymmärtämistä ja hallintaa. He painottavat teknokeskeisten selitettävyyssmenetelmien merkitystä, sillä ne muodostavat lähtökohdan selitysten tuottamiselle monimutkaisissa tekoälyjärjestelmissä. Selittävän tekoälyn kehittämisen lähtökohtana on nykyisten tekoälyjärjestelmien selitettävyyden puute ja tarve vastata tähän ongelmaan. Kehityksessä on tärkeää huomioida kaikkien sidosryhmien tarpeet sekä läpinäkyvyyden eri tasot laajemmassa kontekstissa (Barredo Arrieta ja muut, 2020, s. 86).

Tekoälyjärjestelmien selitettävyyden puute perustellaan usein mallien monimutkaisuudella, kilpailuedun menettämällä ja järjestelmän väärinkäytösten estämisellä (Van Iddekinge ja muut, 2023, s. 671). Barredo Arrieta ja muut (2020, s. 86) toteavat, että selitettävyyden puutteen pääsyyinä ovat tutkimusyhteisön ja liiketoimintasektorin välisen kuilun sekä tuloskeskeisyyden vahvat näkemyserot. Heidän mukaansa liiketoimintasektori suhtautuu vastahakoisesti uusien selitettävyystekniikoiden käyttöönottoon, koska tuloskeskeisyys asetetaan tärkeämmäksi kuin käyttäjien ymmärrys tekoälyn vaikutuksista. Yhteiskunnalliset tekijät ja markkinatoimijat ovat ratkaisevassa asemassa selitettävyyssongelmien ratkaisemisessa. Selitettävyydelle on kuitenkin suurta tarvetta monilla kriittisillä aloilla, kuten pankkialalla, rekrytoinnissa ja terveydenhuollossa (Haresamudram ja muut, 2023, s. 96).

Ei-asiantuntija loppukäyttäjille tekoälyjärjestelmän tuottama selitys on laadullinen selitettävyyssprosessi, jossa pyritään tarjoamaan käyttäjille yksinkertaisilla selitettävyyssmenetelmillä parempaa ymmärrystä taustalla olevista laskennallisista prosesseista ja lopputuloksista (Langer ja muut, 2021, s. 6–7). Ehsan & Riedl (2020) esittävät tapaustutkimuksessaan selittävän tekoälyn suunnittelumenetelmiä, joissa tekoälyjärjestelmä tuottaa havainnollistavia selityksiä ei-teknisille loppukäyttäjille. He havaitsivat, että tekoälyn tekniset edistysaskeleet ja ihmisten teknologinen ymmärtäminen kehittyvät rinnakkain, kun järjestelmä tarjoaa selityksiä käyttäjille. Käyttäjakeskeinen näkökulma on yhtä tärkeä kuin tekninen puoli selitettävyyssmenetelmien kehittämisessä, sillä tekoälyn käyttäjä arvioi järjestelmän

tuottamat selitykset inhimillisten tekijöiden, kuten tunteidensa ja kognitiivisten kykyjensä perusteella (Bernardo & Rosemary, 2023, s. 31).

Inhimilliset tekijät, kuten kausaalisuus, luottamus ja oikeudenmukaisuus, ohjaavat selittävän tekoälyn laadullisia näkökohtia eri sidosryhmien näkökulmasta (Barredo Arrieta ja muut, 2020, s. 86). Barredo Arrieta ja muiden (2020, s. 86–87) korostavat, että inhimilliset tekijät auttavat määrittelemään selitettävyyden tavoitteita laajemmin ja vastaamaan kysymyksiin siitä, mitä varten selitykset annetaan tekoälyjärjestelmissä. Inhimillisten tekijöiden huomioidussa on tärkeää ymmärtää, että algoritmisen toiminnan täydellinen selittäminen ei-asiantuntija käyttäjille on käytännössä mahdotonta ja tarpeetonta, koska laskennalliset prosessit ovat yleensä liian laajoja ja monimutkaisia (Shneiderman, 2020, s. 10).

Tekoälyjärjestelmien laadulliselle selitettävyydelle ei-asiantuntija sidosryhmille on kasvavaa tarvetta, sillä sekä yhteiskunnassa että tiedeyhteisössä on huolta siitä, miten tekoälyjärjestelmät vaikuttavat ihmisten jokapäiväiseen elämään ilman selityksiä perusteluja päätöksentekoprosessista tai lopputuloksista (Shneiderman, 2020, s. 8). Langer ja muut (2021, s. 15–16) korostavat, että selitettävyyden kehittämiseksi on tärkeää painottaa vuorovaikutuksellisia ja yhteiskunnallisia läpinäkyvyyden tasoja, jotta selityksien laadullinen merkitys voidaan ymmärtää ja kehittää ei-asiantuntijoiden näkökulmasta. Myös Ehsan & Riedl (2020, s. 16) painottavat reflektiivisen sosioteknisen lähestymistavan merkitystä selitettävyyden kehittämiseksi, koska se mahdollistaa syvällisemmän ymmärryksen tekoälyjärjestelmien päätöksentekoprosesseista ja niiden vaikutuksista yhteiskuntaan. Sosiotekninen lähestymistapa voi lisätä luottamusta ja hyväksyttävyyttä tekoälyjärjestelmiä kohtaan, koska se huomioi inhimilliset ja eettiset näkökohdat tekoälyn selitettävyyden kehittämiseksi (Bernardo & Rosemary, 2023, s. 6).

Selittävässä tekoälyssä XAI-menetelmät viittaavat tekniikoihin ja lähestymistapoihin, joiden tavoitteena on tehdä tekoälyjärjestelmien toiminnasta läpinäkyvämpää ja

ymmärrettävämpää (Schoonderwoerd ja muut, 2021, s. 4). Barredo Arrieta ja muut (2020, s. 87–88) Barredo Arrieta ja muut (2020, s. 87–88) esittävät, että tekoälymallit voidaan jakaa kahteen kategoriaan, jotka ovat valkoiset ja mustat laatikot. Guidotti ja muut (2019, s. 5) kuvaavat valkoiset laatikot läpinäkyvinä tekoälymalleina, joissa sisäinen toiminta on valmiiksi helpommin ymmärrettävissä ja tulkittavissa. Kun taas Dwivedi ja muut (2023, s. 10) toteavat, että mustan laatikon mallit eivät ole itsestään ymmärrettäviä tai tulkittavia, vaan ne vaativat jälkiselitettäviä analyyseja, jotta niiden toiminta voidaan avata käyttäjille. Mustan laatikon mallit liitetään yleensä jälkiselitettäviin menetelmiin, jotka täydentävät mallien tulkintaa selkeyttämällä niiden päätöksenteon prosessia.

Luonnostaan läpinäkyviä tekoälymalleja ovat esimerkiksi lineaariset regressiomallit, päätöspuut ja yleiset additiiviset mallit (Dwivedi ja muut, 2023, s. 10). Barredo Arrieta ja muut (2020, s. 89–90) korostavat, että osa läpinäkyvistä malleista on valmiiksi simuloitavissa, eli ajateltavissa ihmisten toimesta kokonaisuutena. Kuitenkin osa näistä malleista saattaa monimutkaisuutensa tai laajuutensa vuoksi tarvita XAI-metodeja selitettävyyden parantamiseksi. Tavoitteena on lisätä mallien tulkittavuutta matemaattisten analyysien ja menetelmien avulla, mikä auttaa käyttäjiä ymmärtämään niiden toiminnan perusteita (Barredo Arrieta ja muut, 2020, s. 87).

Bernardo & Rosemary (2023, s. 2) toteavat, että selittävä tekoäly on edistynyt mustan laatikon ongelman ratkaisemisessa, mikä tarkoittaa tekoälyjärjestelmien ja mallien algoritmisen läpinäkyvyyden parantamista. Mustan laatikon ongelmien ratkaisemiseksi kehitetään XAI-menetelmiä, jotka parantavat mustien laatikoiden ennustemallien läpinäkyvyyttä ja tulkittavuutta sekä auttaen ymmärtämään, mitkä tekijät vaikuttavat mallien tekemisiin päätöksiin (Guidotti ja muut, 2019, s. 10). Barredo Arrieta ja muut (2020, s. 93) luokittelevat XAI-menetelmät viiteen pääluokkaan, jotka ovat yksinkertaistamismenetelmät, ominaisuusperusteiset menetelmät, paikallinen selitettävyyden menetelmät, visuaaliseen selitettävyyden menetelmät ja arkkitehtuuriset modifikaatiot. Näillä menetelmillä analysoidaan tekoälymallien eri osia, kuten syötteitä, parametreja tai

laskelmia, mikä parantaa algoritmien läpinäkyvyyttä ja ymmärrettävyyttä kokonaisuudessaan (Barredo Arrieta ja muut, 2020, s. 87).

Yksinkertaistamismenetelmiä käytetään monimutkaisten tekoälymallien yksinkertaistamiseen (Barredo Arrieta ja muut, 2020, s. 87–88). Guidotti ja muiden (2019, s. 29) mukaan tunnetuin yksinkertaistamismenetelmä on LIME (Local Interpretable Model-agnostic Explanations), joka luo yksinkertaisen ja tulkittavan mallin monimutkaisten mallien ennusteiden selittämiseksi esimerkiksi lineaarisen mallin avulla. LIME-menetelmää käytetään myös tekoälymallien paikalliseen selitettävyyteen, jossa pyritään selittämään mallin päätöksiä tietyissä yksittäisissä tapauksissa (Barredo Arrieta ja muut, 2020, s. 92).

SHAP (SHapley Additive exPlanations) on yksi keskeisistä ominaisuusperusteisista XAI-menetelmistä, joka tarjoaa syvällistä ymmärrystä mallien päätöksenteosta (Barredo Arrieta ja muut, 2020, s. 92). Salih ja muut (2024, s. 2) toteavat, että SHAP perustuu peliteorian Shapley-arvoihin ja arvioi, kuinka paljon kukin syöte tai ominaisuus vaikuttaa mallin ennusteeseen. Tämä tehdään laskemalla jokaiselle syötteelle arvo, joka kuvaa sen vaikutusta ennustetulokseen verrattuna mallin keskimääräiseen ennusteeseen (Salih ja muut, 2024, s. 3).

Visuaalisen selitettävyyden tunnettu XAI-menetelmä on Grad-CAM (Gradient-weighted Class Activation Mapping) (Dwivedi ja muut, 2023, s. 13). Barredo Arrieta ja muut (2020, s. 96–97) toteavat, että Grad-CAM on erityisesti suunniteltu konvoluutiomalleille, kuten syväoppiville neuroverkoille, ja se tarjoaa visuaalisesti selitettävän näkymän siihen, mitkä osat kuvasta vaikuttavat mallin ennusteeseen. Grad-CAM luo lämpökarttoja, jotka näyttävät, mitkä alueet syötekokemuksesta vaikuttavat voimakkaimmin mallin päätöksiin, tarjoten käyttäjille visuaalisen ja intuitiivisen tavan ymmärtää, kuinka malli tekee päätöksiä (Dwivedi ja muut, 2023, s. 13).

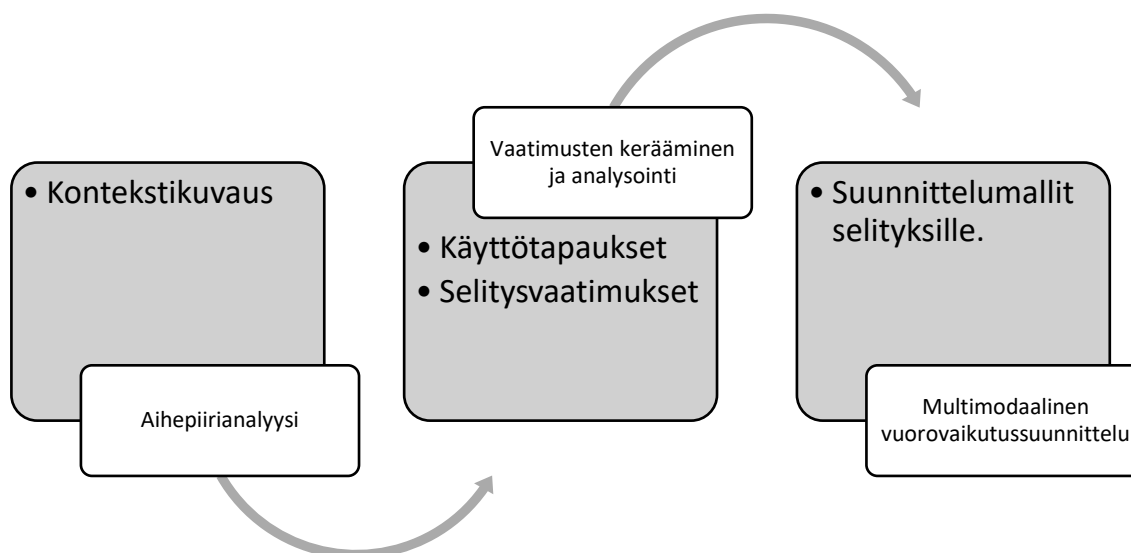
Tunnettu XAI-menetelmä, joka liittyy arkkitehtuurisiin modifikaatioihin on XNNs (Explainable Neural Networks) menetelmä (Guidotti ja muut, 2019, s. 16–17). Barredo Arrieta ja muiden (2020, s. 95–96) toteavat, että nämä menetelmät keskittyvät muokkaamaan neuroverkkojen arkkitehtuuria niin, että malli on luonnostaan läpinäkyvämpi ja sen päätöksenteko on helpommin ymmärrettävää. Tämä tarkoittaa, että malli rakennetaan alusta alkaen siten, että sen toiminta ja päätöksentekoprosessi ovat läpinäkyvämpiä ja intuitiivisempia, mikä helpottaa tulkintaa ja ymmärtämistä (Barredo Arrieta ja muut, 2020, s. 97–98).

3.2.2 Vaatimusten analysointi ja määrittely

XAI-sovellussuunnittelun ensimmäinen vaihe on vaatimusten analysointi, jossa määritellään tekoälysovelluksen käyttötarkoitukset, tekniset vaatimukset toiminnallisuuksille ja ominaisuudet sekä selitettävyysskomponentit (Clement ja muut, 2023, s. 84–85). Clement ja muut (2023, s. 84–85) korostavat, että käyttötarkoituksen ja teknisten vaatimusten analysointi sekä määrittely muodostavat tekoälysovelluksen lähtökohtaisen funktionaalisuuden. Tällä tarkoitetaan ehtoja ja asetuksia, jotka varmistavat, että tekoälysovellus toimii optimoidusti. Selitettävyysskomponenteilla viitataan XAI-menetelmiin, joiden avulla varmistetaan, että tekoälysovelluksen toiminta on läpinäkyvää ja ymmärrettävää loppukäyttäjälle (Schoonderwoerd ja muut, 2021, s. 2).

Schoonderwoerd ja muut (2021, s. 2–3) esittävät DoReMi-käytäntöön perustuvan käyttäjakeskeisen prosessin, jossa kehitetään yleistettäviä XAI-suunnittelumalleja XAI-järjestelmien selitysten esittämiseen. Tässä prosessissa aloitetaan aihepiirianalyysillä, jonka tavoitteena on muodostaa kontekstikuvaus ja konsepti selityksille. Seuraavana vaiheena on vaatimusten kerääminen ja analysointi, jossa muodostetaan käyttötapauksia ja selitysvaatimuksia. Viimeisenä vaiheessa keskitytään multimodaaliseen vuorovaikutussuunnitteluun, jonka tarkoituksena on kehittää suunnittelumallit selityksille. XAI-suunnittelumallien kehittämisessä keskeistä on käyttäjien tarpeiden syvälinen analysointi ja kontekstisidonnaisten ratkaisujen luominen

laadullisen prosessin kautta (Langer ja muut, 2021, s. 15). Alla oleva kuvio 9 esittää DoReMi-käytäntöön perustuvan prosessin ja suunnittelumallin.



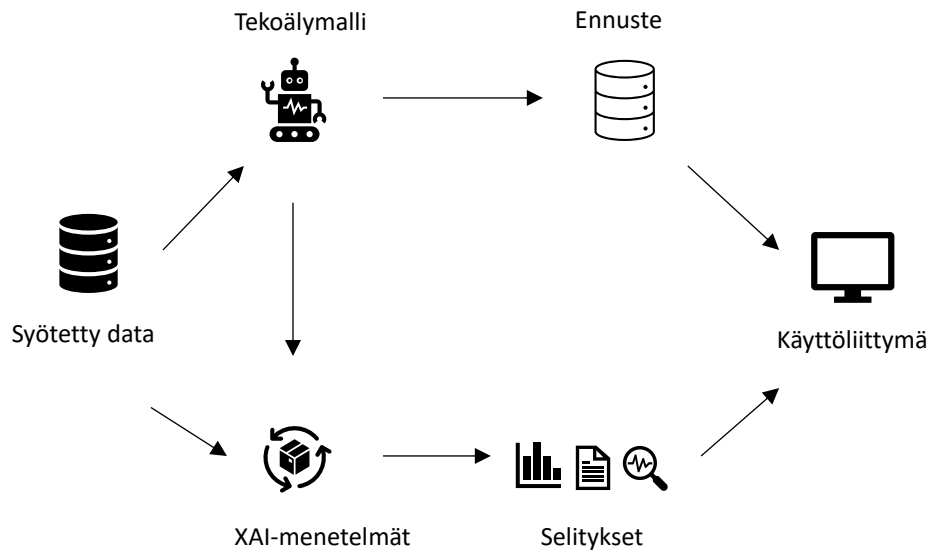
Kuvio 9. DoReMi-suunnittelumalli (Schoonderwoerd ja muut, 2021, s. 2).

Vaatimusten keräämiseen ja analysointiin on useita menetelmiä, mutta tärkeintä on tunnistaa ja määritellä, mitkä ovat ne vaatimukset, joita käyttäjä asettaa järjestelmän antamille selityksille ja miten selitykset tulisi mukauttaa sovellusskenaarion kontekstiin (Schoonderwoerd ja muut, 2021, s. 3). Clement ja muut (2023, s. 85) toteavat, että vaatimusten keräämisen ja analysoinnin jälkeen saadaan tarkempi käsitys siitä, millaista tietoa käyttäjät odottavat järjestelmän selityksiltä. Vaatimusten keräämiseen ja analysointiin tehokas menetelmä on pyytää käyttäjiä antamaan itse selityksiä käyttötapausten perusteella. Tämä lähestymistapa tuottaa alustavaa tietoa heidän odotuksistaan ja auttaa tunnistamaan lisävaatimuksia järjestelmän selityksille (Schoonderwoerd ja muut, 2021, s. 3).

3.2.3 Multimodaalinen vuorovaikutussuunnittelu

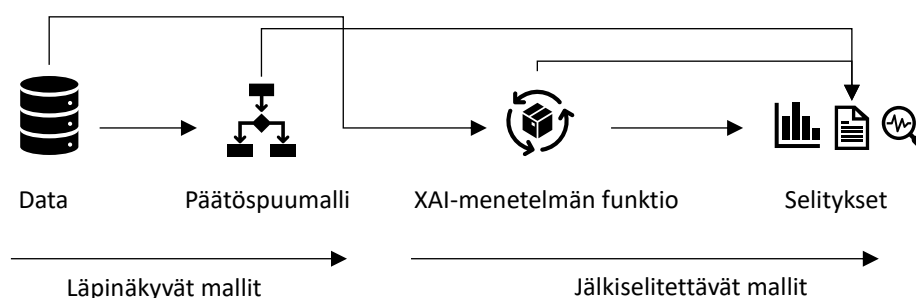
XAI-suunnittelussa multimodaalinen vuorovaikutussuunnittelu jakaantuu kahteen osaan, selitettävyyssuunnitteluun ja selittävän käyttöliittymän suunnitteluun (Clement ja muut, 2023, s. 85–86). Schoonderwoerd ja muut (2021, s. 4) määrittelevät, että selitettävyyssuunnittelun tavoitteena on kehittää selitysmalli, joka kuvaa prosessin, miten tekoälyjärjestelmän on tarkoitus tuottaa selityksiä ja mitä tukevia XAI-menetelmiä tähän tarvitaan. Selittävän käyttöliittymän suunnittelussa selitysmallista muodostetaan konkreettinen esitys, kuten prototyyppi, joka sisältää vuorovaikutukselliset selitettävyyteen liittyvät ominaisuudet ja toiminnallisuudet (Clement ja muut, 2023, s. 86).

Selitettävyyssuunnittelussa oleellisinta on tunnistaa XAI-menetelmien funktionaalinen toiminta ja yhdistää ne käytettäviin tekoälymalleihin (Clement ja muut, 2023, s. 86). Schoonderwoerd ja muut (2021, s. 3) korostavat, että selityksissä on tärkeää tunnistaa sekä tulostettavan tiedon tyyppi että tapa, jolla tieto esitetään käyttäjälle. Nämä valinnat määräytyvät selitettävyystavoitteiden sekä XAI-menetelmien ominaisuuksien ja toiminnallisuuksien mukaan, kuten esimerkiksi ominaisuusperusteisessa selitysmallissa tai yksinkertaistamismenetelmissä, jotka korostavat päätöksiin vaikuttavia tärkeimpiä tekijöitä (Barredo Arrieta ja muut, 2020, s. 87). Alla oleva kuvio 10 havainnollistaa tekoälymallin ennusteen tuottamisprosessin syötetystä datasta käyttöliittymään asti sekä XAI-menetelmien roolin ennusteiden selittämisessä käyttäjille. Prosessi muodostaa syklin, jossa tavoitteena on parantaa käyttäjän ymmärrystä tekoälymallin toiminnasta ja tukea vastuullista päätöksentekoa.



Kuvio 10. Ennusteen ja selittämisen tuottamisen prosessi (Clement ja muut, 2023, s. 81).

Vilone & Longo (2021, s. 616) toteavat, että selitysmallien luomisessa lähestytään kahdesta toisistaan eroavasta näkökulmasta, joita ovat luonnostaan läpinäkyvät mallit ja jälkiselitettävät mallit. XAI-menetelmiä käytetään pääsääntöisesti jälkiselitettävissä malleissa, mutta niitä voidaan hyödyntää myös läpinäkyvissä malleissa, jos tarvitaan lisäselvityksiä tai syvällisempää tietoa (Barredo Arrieta ja muut, 2020, s. 87–88). Vilone & Longo, 2021 (s. 616) korostavat, että läpinäkyvissä tekoälymalleissa selitysten konstruointi voidaan suunnitella jo tekoälyjärjestelmän rakennusvaiheessa, sillä päätöksentekoprosessit ovat etukäteen tulkittavissa. Jälkiselitettävissä malleissa sen sijaan XAI-menetelmillä pyritään tuottamaan selittävää informaatiota jälkikäteen, kun järjestelmä on jo luotu. Tämän avulla käyttäjät voivat ymmärtää järjestelmän päätöksiä, vaikka mallin sisäinen toiminta ei olisikaan suoraan havaittavissa (Barredo Arrieta ja muut, 2020, s. 88). Alla oleva kuvio 11 havainnollistaa tekoälymallien selitettävyyden lähestymistavat jaoteltuina läpinäkyviin ja jälkiselitettäviin malleihin. Kuviossa esitetään datan käsittely päätöspuumallin kautta sekä XAI-menetelmien rooli selitysten tuottamisessa jälkiselitettäville malleille.



Kuvio 11. Läpinäkyvien ja jälkiselitettävien mallien prosessi (Vilone & Longo, 2021, s. 616).

Selitysten konstruoiminen tarkoittaa prosessia, jossa tekoälyjärjestelmän tuottaman informaation perusteella muodostetaan johdonmukaisia ja ymmärrettäviä selityksiä mallin päätöksenteosta sekä tuloksista (Clement ja muut, 2023, s. 93–94). Chromik & Butz (2021, s. 2) määrittelevät, että kun selityksiin lisätään ominaisuuksia ja toiminnallisuuksia, jotka mahdollistavat käyttäjäryhmien vuorovaikutuksen selitysten kanssa, kyseessä on selittävä käyttöliittymä. Selittävässä käyttöliittymässä tiedon esittäminen ja käyttäjäinteraktiot keskittyvät käyttöliittymiin, joissa korostuvat ominaisuuksien vaikutukset, riippuvuusselitykset, sääntöpohjaiset lähestymistavat, ominaisuuksien konteksti sekä tekoälymallien suorituskyky (Cálem ja muut, 2024, s. 11).

Selittävän tekoälyn käyttöliittymän keskeisiä osa-alueita ovat viisi tunnistettua käyttöliittymäklusteria, joita ovat visuaalinen analytiikka, interaktiivinen analytiikka, digitaaliset verkostot, kuvien visualisointi ja luonnollisen kielen vuorovaikutus (Cálem ja muut, 2024, s. 11). Cálem ja muut (2024) korostavat, että nämä klusterit tarjoavat eri tavoin tukea käyttäjän ymmärrykselle ja päätöksenteolle. Visuaalinen analytiikka keskittyy tietojen graafiseen esittämiseen, kun taas interaktiivinen analytiikka mahdollistaa käyttäjän aktiivisen osallistumisen tiedon tutkimiseen. Digitaaliset verkostot havainnollistavat suhteita ja yhteyksiä tietomallissa, ja kuvien visualisointi auttaa konkretisoimaan monimutkaisia tietorakenteita. Luonnollisen kielen

vuorovaikutus puolestaan tarjoaa käyttäjille mahdollisuuden kommunikoida tekoälyn kanssa helposti ymmärrettävällä kielellä. Ihmisen ja selittävän tekoälyn käyttöliittymän tavoitteena on hyödyntää operationaalisia ja selitettäviä ominaisuuksia tukemaan tavoitteiden saavuttamista (Chromik & Butz, 2021, s. 5).

3.2.4 Vuorovaikutukselliset selitettävyyslähestymistavat

Chromik & Butz (2021, s. 2) määrittelevät, että tyypillisiä vuorovaikutuskonsepteja voidaan muokata selittävän tekoälyn käyttöliittymiin keskittymällä siihen, miten nämä konseptit integroidaan tekoälyn selitystoimintoihin käyttäjäkokemuksen ja käytettävyyden parantamiseksi. Vuorovaikutuskonseptit sisältävät tiedon välittämisen, dialogin, kontrollin, kokemuksen, optimaalisen käyttäytymisen, työkalun käytön ja kehollisen toiminnan (Chromik & Butz, 2021, s. 14). Cálem ja muut (2024, s. 7) korostavat, että näiden konseptien käytännön soveltaminen voi paljastaa uusia mahdollisuuksia ja ongelmia selittävien käyttöliittymien kehittämisessä.

Vuorovaikutus tiedon välittämisenä keskittyy informaation yksisuuntaiseen esittämiseen käyttäjälle (Chromik & Butz, 2021, s. 6). Chromikin & Butzin (2021, s. 6) mukaan selittävän tekoälyn yhteydessä vuorovaikutus tarkoittaa kuitenkin yhteistoimintaa, jossa käyttäjä ja tekoälyjärjestelmä täydentävät toisiaan selitysten avulla. Ehsan ja muut (2019) esittävät tekniikan, jossa käyttäjän toiminta selitetään reaaliaikaisesti tekoälyjärjestelmän toiminnassa, parantaen käyttäjän ymmärrystä vaikutuksista. Binns ja muut (2018) tutkivat, miten erilaiset perustelevaan logiikkaan pohjautuvat selitystyyppit vaikuttavat käyttäjien oikeudenmukaisuuskäsityksiin eri sovelluskenaarioissa. Tiedon välittämisessä korostuvat selitysten tarkkuus ja eri selitystyylien kyky esittää ymmärrettävää tietoa tekoälyn toiminnasta (Chromik & Butz, 2021, s. 6).

Vuorovaikutus dialogina kuvataan tietokoneen ja ihmisen välisten syötteiden ja palautteiden sykliä (Chromik & Butz, 2021, s. 7). Koivunen ja muut (2022, s. 488) toteavat, että ihmisen ja tekoälyjärjestelmän välistä dialogia ilmentävät keskustelevat

käyttöliittymät, kuten chattibotit. Selittävän tekoälyn yhteydessä dialogi tapahtuu vaiheittain ja on käyttäjän ohjaamaa toimintaa, jossa korostuvat selitysten luonnollisuus ja saavutettavuus (Chromik & Butz, 2021, s. 7). Wibergin & Stolterman Bergqvistin (2023, s. 2 287) mukaan keskustelevat käyttöliittymät mahdollistavat vuorovaikutuksen, jossa käyttäjä voi pyytää tarkempia tietoja tai tiettyä esitystapaa tekoälypohjaisten päätöksentekoprosessien suhteen. Tekoäly voi mukautua käyttäjän tarpeisiin automaattisesti.

Kontrolli vuorovaikutuksena viittaa käyttäjän ja tekoälyjärjestelmän kykyyn saavuttaa tavoitetila vakaasti ja nopeasti (Chromik & Butz, 2021, s. 8). Chromikin & Butzin (2021, s. 8) mukaan tämä tarkoittaa selittävässä tekoälyssä sitä, että käyttäjälle annetaan tietoa ymmärrysprosessin tukemiseksi, mutta samalla mahdollistetaan hänen vaikutusmahdollisuutensa tekoälyyn ja sen ohjattuun käyttäytymiseen. Selittävissä käyttöliittymissä kontrolli on ohjaava vuorovaikutustyyli, jossa käyttäjät voivat manipuloida esimerkiksi syöttämäänsä dataa tai muuttaa datan luokitteluja (Cálem ja muut, 2024, s. 7). Cálemin ja muiden (2024) mukaan tämä voi perustua sääntöpohjaiseen tekoälyjärjestelmään, joka selittää, miksi tietty sääntö aktivoitui virhetilanteissa.

Vuorovaikutus kokemuksena käsittää ihmisen odotuksia tietokoneita kohtaan (Chromik & Butz, 2021, s. 9). Chromik & Butz (2021, s. 9) toteavat, että selittävän tekoälyn kontekstissa tämä tarkoittaa käyttäjien odotusten ja mieltymysten hallintaa, jossa keskitytään erityisesti luottamuksen, tyytyväisyyden ja vakuuttavuuden tavoitteisiin. Monimuotoiset selitykset, kuten graafiset ja interaktiiviset elementit, voivat rikastuttaa käyttäjäkokemusta merkittävästi, mutta niiden käyttö edellyttää huolellista suunnittelua, jotta monimutkaisuus ja vuorovaikutteisuus pysyvät tasapainossa (Cálem ja muut, 2024, s. 11).

Optimaalinen käyttäytyminen vuorovaikutuksena keskittyy käyttäjän käyttäytymisen sopeuttamiseen siten, että se tukee paremmin heidän tehtäviään ja tavoitteita (Chromik

& Butz, 2021, s. 10). Chromikin & Butzin (2021, s. 10) mukaan selittävän tekoälyn kontekstissa optimaalinen käyttäytyminen vuorovaikutuksen tavoitteena on ohjata käyttäjiä heitä tyydyttävälle tasolle ymmärryksessä, keskittyen tarjoamaan selityksiä, jotka kouluttavat ihmisiä paremmin vuorovaikutukseen tekoälyn kanssa. He toteavat, että käyttäjät voivat mahdollisesti kohdata esimerkiksi virheellisiä tekoälyjärjestelmiä tai kokea niistä väärinkäsityksiä kognitiivisten vinoumien takia. Optimaalinen käyttäytyminen vuorovaikutuksena tarkoittaa suunnittelijoille käyttäjän ja tekoälyjärjestelmään liittyvien vuorovaikutuksellisten rajoitusten hallintaa paremmaksi käyttäjäkohtaisesti (Cálem ja muut, 2024, s. 7).

Vuorovaikutus työkalun käytön kautta keskittyy siihen, että tietokoneita hyödynnetään laajentamaan käyttäjän kykyjä (Chromik & Butz, 2021, s. 12). Chromikin & Butzin (2021, s. 12) mukaan selittävän tekoälyn kontekstissa tätä lähestymistapaa käytetään oppimisen välineenä, jossa selitysten avulla autetaan käyttäjiä oivaltamaan uusia asioita tietyistä aihealueista. Vuorovaikutustyyli näkyy erityisesti virtuaalisessa ja lisätyssä todellisuudessa, joissa simuloitujen ja pelillisten elementtien yhdistetään osaksi järjestelmiä (Pillai ja muut, 2022, s. 7).

Vuorovaikutus kehollisena toimintana keskittyy käyttäjän ja tietokoneen väliseen symbioottiseen suhteeseen, jossa molemmat osapuolet tekevät yhteistyötä päätöksenteossa (Chromik & Butz, 2021, s. 13). Chromik & Butz (2021, s. 13) toteavat, että tämä vuorovaikutuksen konsepti perustuu ihmisen ja tekoälyn yhteistoimintaan, jossa molemmat osapuolet vaikuttavat toistensa toimintaan reaaliajassa ja oppivat toisiltaan saavuttaakseen yhteisen tavoitteen tehokkaammin. Tämä vuorovaikutustyyli sisältää mekanismin, joka mahdollistaa autonomisen järjestelmän havaitsemaan mallieroja itsensä ja ihmiskumppanin välillä, päättämään erimielisyyden lähteen, arvioimaan virheen mahdollisia seurauksia ja lopulta tarjoamaan ihmiselle ymmärrettävää palautetta mallin korjaamiseksi (Tabrez ja muut, 2019, s. 256).

4 Tutkimusmenetelmä- ja aineisto

4.1 Tutkimusmenetelmä

Tämä tutkimus on laadullinen eli kvalitatiivinen tutkimus. Kuten Tuomi & Sarajärvi (2018, s. 18) toteavat, kvalitatiivisen tutkimuksen päämääränä on tuottaa syvällistä tietoa tutkittavasta ilmiöstä, keräämällä aineistoa luonnollisista tilanteista. Induktiivinen tutkimusprosessi etenee yksittäisistä havainnoista kohti yleisempiä johtopäätöksiä, huomioiden samalla useita erilaisia tekijöitä, jotka voivat vaikuttaa tutkimustuloksiin. Tutkimuksen edetessä luokat ja teemat muodostuvat havaintojen perusteella. Kvalitatiivisessa tutkimuksessa pyritään ymmärtämään tutkittavan ilmiön moninaisuutta ja merkityksiä, korostaen tulkintaa ja aineiston kontekstuaalisuutta, eikä niinkään yleistettävyyttä, kuten kvantitatiivisessa tutkimuksessa (Alasuutari, 2011, s. 72).

Tietojärjestelmätieteessä on useita erilaisia tutkimusmenetelmiä, erityisesti laadulliseen tutkimukseen liittyviä. Goldkuhlin (2019, s. 575) mukaan laadullinen lähestyminen ei itsessään yleensä ole riittävä, sillä tietojärjestelmätieteessä korostuvat myös aineistojen perusteella luodut ratkaisut, kuten digitaaliset suunnittelumallit ja artefaktit. Tietojärjestelmätutkimuksen kenttä hyödyntää monia erilaisia lähestymistapoja, kuten tapaustutkimusta, toimintatutkimusta ja suunnittelutiedettä. Näiden menetelmien avulla voidaan tutkia ja kehittää käytännön ratkaisuja, malleja ja prosesseja, jotka ovat keskeisiä tietojärjestelmien ymmärtämisessä ja kehittämisessä. Konstruktiivinen tutkimusote sisältää edellä mainittujen lähestymistapojen erityispiirteitä, mutta siinä keskitytään myös ongelmalähtöiseen tutkimusprosessiin, jossa ydin on ongelmien tunnistaminen ja ratkaisujen kehittäminen (Rolin ja muut, 2006, s. 112).

Tämä tutkimus perustuu konstruktiiviseen lähestymistapaan, jossa painotetaan uusien ratkaisujen luomista ja kehittämistä. Rolin ja muut (2006, s. 113) kuvaavat konstruktiivisen tutkimusotteen ydinpiirteet seuraavasti, ongelman ja ratkaisun käytännöllinen merkitys, ratkaisun toimivuus käytännössä, yhteys aikaisempaan teoriaan ja tutkimuksen teoreettinen kontribuutio. Näitä ydinpiirteitä voidaan tarkastella

sekä interpretatiivisesta että pragmaattisesta näkökulmasta (Goldkuhl, 2012, s. 143). Goldkuhl (2012, s. 144) huomauttaa, että interpretatiiviset piirteet korostavat syvällistä ymmärrystä ja teoreettista yhteyttä, kun taas pragmaattiset piirteet keskittyvät ratkaisujen käytännön merkitykseen ja toimivuuteen.

Kasanen ja muut (1993) toteavat, että konstruktiiivinen tutkimusprosessi käynnistyy todellisessa elämässä havaitun ongelman tunnistamisesta, joka on sekä käytännöllisesti merkityksellinen että teoreettisesti kiinnostava. Tutkimuksessa pyritään kehittämään ratkaisu, joka vastaa tähän ongelmaan, ja ratkaisu voi toimia mallina käytännön sovelluksille. Tämän jälkeen tutkijan tulee hankkia laaja ymmärrys aiheesta ennen varsinaisen ratkaisun kehittämistä. Innovaatio, eli ratkaisun luominen, on olennainen vaihe, jota seuraa ratkaisun toimivuuden todentaminen. Lopuksi tutkija osoittaa ratkaisun teoreettiset kytkökset ja tutkimuksellisen merkityksen, sekä arvioi ratkaisun sovellettavuutta eri konteksteissa.

4.2 Haastateltavat

Haastateltavien valinnassa on tärkeää ottaa huomioon tutkimuksen kohde. Vilka (2021, s. 108) toteaa, että haastateltavat tulisi valita tutkimusongelman mukaisesti, painottaen heidän asiantuntemustaan tai kokemustaan tutkimusaiheesta tai käsiteltävästä ilmiöstä. Tässä tutkimuksessa oli alustavasti ongelmana valita haastateltavia, sillä Hunkenschroerin & Luetgen (2022, s. 996) mukaan tekoälypohjaisten rekrytointijärjestelmien käyttö sekä tiedottaminen rekrytointiprosessien yksityiskohdista työnhakijoille on organisaatioissa usein vähäistä tai lähes olematonta. Tämä havaittiin myös käytännössä, sillä monet potentiaaliset haastateltavat, eli työnhakijat, eivät osanneet varmuudella sanoa, olivatko olleet tekoälypohjaisten rekrytointijärjestelmien käytön kohteena. Koska haastateltavien omakohtaisista kokemuksista tekoälypohjaisesta rekrytointijärjestelmästä ei voitu saada varmuutta, niitä ei voitu käyttää täysin varmistettuna valintakriteerinä, jonka seurauksena tutkimus ohjautui heidän käsitysten ja tulkintojen tutkimiseen aiheesta.

Käsitysten tutkimiseen voidaan valita haastateltavia useammalla tavalla. (Vilka, 2021, s. 109) toteaa, että käsitysten ei välttämättä tarvitse olla muotoutuneet omakohtaisen kokemuksen avulla, vaan niihin vaikuttavat esimerkiksi yhteisöjen perinteet, tyypillinen ajattelutapa tai yleinen puhe. Haastateltavilla saattaa olla paljon käsityksiä tekoälypohjaisista rekryointijärjestelmistä ja niiden läpinäkyvyydestä tai päätöksien selitettävyydestä ilman varmuutta kokemuksesta. Tämän vuoksi haastateltavien vallinnassa on olennaista tunnistaa, että käsitykset voivat perustua laajempiin yhteiskunnallisiin diskursseihin tai ammatillisiin keskusteluihin, eivätkä omakohtaiseen kokemukseen. Näin ollen tutkimus voi tuottaa arvokasta tietoa siitä, miten tekoälypohjaisten rekryointijärjestelmien käyttöä ja läpinäkyvyyttä käsitetään, vaikka haastateltavat eivät olisi itse suoraan osallistuneet tekoälyn ohjaamaan rekryointiprosessiin. Tämä lähestymistapa mahdollistaa monipuolisemman näkökulman järjestelmien vaikutuksiin ja päätöksen selitettävyydestä vallitseviin käsityksiin.

Tutkimuksen haastateltavien valintakriteerit edellyttävät, että haastateltavilla on vähintään kerran vuoden sisällä kokemusta työhausta LinkedInin rekryointijärjestelmän käytöstä. Black & van Esch (2020) toteavat, että kyseisellä yrityksellä on monipuolinen valikoima tekoälypohjaisia rekryointityökaluja, mikä viittaa siihen, että haastateltavat ovat todennäköisesti olleet tekoälypohjaisen rekryoinnin kohteena, vaikka eivät sitä tiedostaneetkaan. Valintakriteereihin kuuluu myös vähintään 25 vuoden ikä, jotta haastateltavilla on riittävästi elämäkokemusta arvioida rekryointiprosessia ja tekoälyjärjestelmien vaikutuksia. Tavoitteena on saada kattava ja monipuolinen näkemys siitä, miten erilaisista taustoista tulevat työnhakijat käsittelevät ja tulkitsevat tekoälyn käyttöä rekryoinnissa, mikä voi johdattaa uusiin ideoihin rekryointijärjestelmien läpinäkyvyyden edistämiseen selitettävyysoimintojen kautta.

4.3 Aineisto

Tutkimuksen aineisto kerättiin teemahaastatteluiden muodossa. Vilka (2021, s. 99) korostaa, että teemahaastattelumenetelmällä on mahdollista tutkia syvällisesti yksilön

kokemuksia ja ajatusrakenteita. Tuomi & Sarajärvi (2018, s. 65) toteavat, että teemahaastattelu muistuttaa avoimuudessaan puolistrukturoitua haastattelua, mutta etenee ennalta määriteltyjen teemojen ja niihin liittyvien kysymysten mukaisesti. Tavoitteena on löytää merkityksellisiä vastauksia tutkimusongelmaan, minkä vuoksi teemat perustuvat tutkimuksen viitekehykseen (Puusa ja muut, 2020, s. 107). Valitsin teemahaastattelun menetelmäksi, koska sen avoimuus ja joustavuus mahdollistavat tutkittavien näkemysten ja kokemusten syvällisen tarkastelun. Tämä menetelmä tarjoaa tilaa sekä ennalta valittujen teemojen käsittelyyn että uusien, tutkimuksen kannalta merkityksellisten näkökulmien esiin nousemiselle.

Tässä tutkimuksessa käsitellään tutkimuskysymysten mukaisesti tekoälypohjaisiin rekrytointijärjestelmiin liittyviä odotuksia vuorovaikutuksellisesta läpinäkyvyydestä. Tuomen & Sarajärven (2018, s. 65) mukaan teemahaastattelussa vuorovaikutus haastattelijan ja haastateltavien välillä etenee vapaasti. Haastattelija voi ohjata keskustelua ennalta laadittujen kysymysten pohjalta, mutta tavoitteena on myös antaa tilaa haastateltavien omille näkemyksille. Yksi merkittävistä eduista on se, että haastattelun kulkua voidaan joustavasti mukauttaa haastateltavien mukaan (Puusa ja muut, 2020, s. 107). Kysymysrunko muodostettiin työnhakijoita ajatellen, sillä haluttiin varmistaa, että heidän odotuksensa ja kokemuksensa tekoälypohjaisten rekrytointijärjestelmien selitettävyydestä tulevat esille. Tavoitteena oli paitsi syventää aiemman tutkimuksen pohjalta nousevia teemoja, myös tuoda esiin uusia näkökulmia, jotka liittyvät vuorovaikutukselliseen läpinäkyvyyteen rekrytointiprosessissa.

Konstruktiiviselle tutkimukselle on tyypillistä toteuttaa tutkimus yhteistyössä kohdeorganisaation kanssa, mutta tässä tutkimuksessa painopiste on enemmän teoreettisessa tarkastelussa kuin käytännön yhteistyössä. Tämän vuoksi aineistonkeruu tapahtui ilman suoraa sidettä tiettyyn organisaatioon, mikä mahdollistaa laajemman näkökulman tutkimusongelmaan. Tämä valinta tukee tutkimuksen tavoitteita tuottaen yleisemmin soveltavia johtopäätöksiä, joita ei rajoita tietyn organisaation erityispiirteet.

Tämän tutkimuksen aineiston muodosti kuuden henkilön haastattelut. Tarkoituksena oli haastatella työnhakijoita tekoälypohjaisiin rekrytointijärjestelmiin liittyen, jotta saataisiin mahdollisimman kattava näkemys heidän odotuksistaan ja kokemuksistaan liittyen selitettävyystietoiseen vuorovaikutukseen. Haastateltavien määrä valikoitui sen perusteella, että se mahdollisti riittävän monipuolisen aineiston, mutta toisaalta säilytti mahdollisuuden syvälliseen keskusteluun jokaisen haastateltavan kanssa. Haastattelut pidettiin etänä. Haastattelujen kesto vaihteli 45 minuutista tuntiin, riippuen haastateltavan näkemyksistä ja siitä, kuinka paljon uusia teemoja nousi esille keskustelun aikana.

4.4 Aineiston analyysi

Haastatteluista kerätty aineisto analysoitiin käyttäen sisällönanalyysiä. Puusan ja muiden (2020, s. 176) mukaan sisällönanalyysi toimii laadullisessa tutkimuksessa systemaattisena menetelmänä, jonka kautta aineisto voidaan tiivistää ja jäsentää empiiristä tulkintaa varten. Se ei ole yksi ainoa menetelmä, vaan pikemminkin joustava viitekehys, jonka avulla aineistoa voidaan tarkastella monipuolisesti. Tuomi & Sarajärvi (2018, s. 83) erittelevät sisällönanalyysille kolme erilaista lähestymistapaa: aineistolähtöinen, teoriaohjaava ja teorialähtöinen analyysi. Kaikissa näissä pyritään tulkitsemaan tekstin merkitysisältöä, mutta aineiston hankinnassa, analyysissä sekä raportoinnissa on vaihtelua lähestymistapojen välillä (Tuomi & Sarajärvi, 2018, s. 83).

Tässä tutkimuksessa käytettiin aineistolähtöistä sisällönanalyysiä. Tuomen & Sarajärven (2018, s. 80) mukaan aineistolähtöisessä analyysissä pyritään rakentamaan teoreettinen kokonaisuus valitsemalla analyysiyksiköt aineistosta tutkimuksen tavoitteiden ja asetettujen kysymysten perusteella. Analyysiyksiköt eivät ole etukäteen sovittuja, eikä teorian merkitys ole analyysiyksiköiden ohjaavana tekijänä keskeinen, vaan analyysi etenee aineiston pohjalta nousevien teemojen mukaisesti. Aineistolähtöisen analyysin painopiste on aineiston sisältämissä merkityksissä, ja teorian merkitys on tutkimuksessa ohjaavana viitekehystenä, mutta ei rajoittava (Vilka, 2021, s. 134).

Aineistolähtöisessä induktiivisessa analyysissä analyysiyksikkö, joka voi olla sana, lause tai ajatuskokonaisuus, määritellään ennen varsinaisen analyysin aloittamista (Tuomi & Sarajärvi, 2018, s. 90). Tämän jälkeen alkaa Tuomen & Sarajärven (2018, s. 92) mukaan aineistolähtöisen analyysin ensimmäinen vaihe, joka on redusointi eli pelkistäminen, jossa alkuperäisilmauksista tehdään pelkistetyt ilmaukset. Tässä vaiheessa aineisto tiivistetään poistamalla epäolennainen tieto ja säilyttämällä analyysin kannalta keskeiset osat. Aineiston pelkistämistä ohjaa tutkimustehtävä, joka auttaa rajaamaan aineistoa samankaltaisten ja erilaisten ilmaisujen joukoksi, koodaamalla tai merkitsemällä ne tunnisteilla (Puusa ja muut, 2020, s. 147).

Aineistolähtöisen analyysin seuraavana vaiheena on klusterointi eli ryhmittely, jossa käydään aineisto tarkemmin läpi ja etsitään samankaltaisuuksia sekä eroavaisuuksia käsitteiden suhteen (Tuomi & Sarajärvi, 2018, s. 92). Tuomen & Sarajärven (2018, s. 92) mukaan tässä vaiheessa samaa ilmiötä kuvaavat käsitteet ryhmitellään ja järjestetään eri luokkiin, joista muodostuvat alaluokat, käyttäen luokitteluyksiköinä ilmiön ominaisuuksia, käsityksiä tai piirteitä. Tätä vaihetta voidaan kutsua myös teemoitteluksi, jossa määriteltyjä koodeja tai tunnisteita hyödynnetään teemojen muodostamisessa (Puusa ja muut, 2020, s. 147). Puusa ja muut (2020, s. 145) toteavat, että tässä vaiheessa tutkija hahmottaa teemojen sisäisiä rakenteita ja niiden suhdetta toisiinsa, mikä syventää ymmärrystä analysoitavasta aineistosta ja sen merkityksestä.

Aineistolähtöisen analyysin viimeisessä vaiheessa tehdään abstrahointi eli käsitteellistäminen (Tuomi & Sarajärvi, 2018, s. 93). Tuomi & Sarajärvi (2018, s. 93) korostavat, että tässä vaiheessa erotetaan tutkimuksen kannalta oleellinen tieto, ja sen perusteella kehitetään teoreettisia käsitteitä. Tämä prosessi sisältää siirtymisen alkuperäisistä kielellisistä ilmauksista kohti yleistettyjä käsitteitä ja johtopäätöksiä, mikä tarkoittaa yhdistämällä luokituksia sisällön näkökulmasta. Tarkoituksena on nostaa aineistosta esiin merkitysten muodostamia kokonaisuuksia teorian ohjaamana, jossa samalla välittyy tutkijan päättelyketju analyysin suhteen (Puusa ja muut, 2020, s. 149).

Aineistolähtöisessä analyysissä korostuu tutkimusaineiston perusteellinen käsittely, mikä tarkoittaa, että analyysin käsitteet ja johtopäätökset sidotaan teoriaan. Vilkan, 2021 (s. 134) mukaan aineistolähtöinen analyysi jatkuu aineiston analyysin jälkeen tulkintavaiheeseen, jossa analyysin löydökset kytketään teoreettiseen viitekehykseen. Tuomi & Sarajärvi (2018, s. 81) korostavat, että teoria tukee analyysia, mutta aineistosta nousevat teemat, ilmiöt, käsitteet ja johtopäätökset muodostavat sen perustan. Teoriaa hyödynnetään löydösten jäsentämisessä sekä niiden ymmärtämisessä laajemmassa kontekstissa (Puusa ja muut, 2020, s. 151).

Aineistoanalyysin jälkeen tutkimuksessa koottiin tulokset tekstiksi ja muodostettiin niistä johtopäätöksiä viitaten tutkimuskysymyksiin. Tämä tutkimus rakentuu konstruktiviselle tutkimukselle, joten tavoitteena oli löytää merkityksellisiä oivalluksia tekoälypohjaisten rekrytointijärjestelmien vuorovaikutuksellisen läpinäkyvyyden kehittämiseksi sekä muodostaa raportti tämän perusteella. Raportin tarkoituksena on tarjota informaatiota selitettävyydestä suunnitteluun tekoälypohjaisissa rekrytointijärjestelmissä, jossa huomioidaan käyttäjäryhmänä päätösten kohteena olevat työnhakijat.

5 Tutkimuksen tulokset

Tässä luvussa esitellään tutkimuksen keskeiset tulokset, jotka pohjautuvat työnhakijoiden näkemyksiin tekoälypohjaisten rekrytointijärjestelmien selitettävyydestä ja vuorovaikutuksellisuudesta. Tulokset on jäsennelty aineistosta nousseiden teemojen mukaisesti, ja niitä tarkastellaan suhteessa tutkimuskysymyksiin. Ensimmäinen tutkimuskysymys pyrki selvittämään, millaisia selitettävyytlähestymistapoja työnhakijat odottavat tekoälypohjaisten rekrytointijärjestelmien tarjoavan. Toisen tutkimuskysymyksen tavoitteena oli tarkastella, millä tavoin tekoälyjärjestelmien päätöksiä ja käsittelemiä tietoja tulisi esittää työnhakijoille rekrytointiprosessin aikana vuorovaikutuksellisella ja ymmärrettävällä tavalla.

Aineison analyysin perusteella nousi esiin viisi keskeistä teemaa, jotka liittyvät selitettävyyden tarpeisiin, vuorovaikutteisuuden merkitykseen, palautteen saamiseen sekä tekoälyn oikeudenmukaisuuteen ja läpinäkyvyyteen. Työnhakijat toivoivat erityisesti selkeitä ja helposti ymmärrettäviä selityksiä, joissa otetaan huomioon heidän henkilökohtaiset tarpeensa ja hakemuksensa ominaisuudet. Vuorovaikutteisuus ja mahdollisuus kysyä tarkennuksia sekä vertailutietoja muihin hakijoiden nähden olisivat myös keskeisiä toiveita. Seuraavissa alaluvuissa esitetään tarkemmin nämä teemat ja analyysistä nousseet keskeiset havainnot. Haastateltavien sitaattit merkitään tunnisteilla H1-H6 tunnisteilla.

5.1 Työnhakijoiden odotukset tekoälypohjaisten rekrytointijärjestelmien selitettävyydestä

Tutkimuksen ensimmäinen keskeinen johdattelava havainto liittyy työnhakijoiden odotuksiin siitä, miten tekoälypohjaisten rekrytointijärjestelmän päätöksiä ja toimintoja tulisi selittää. Tuloksissa kävi ilmi, että työnhakijat toivovat ennen kaikkea selkeitä, helposti ymmärrettäviä ja konkreettisia selityksiä tekoälyjärjestelmän tuottamien päätösten suhteen. Useat haastateltavat painottavat, että selitysten tulisi olla mahdollisimman yksinkertaisia, joissa vältetään syvällisen teknisen termistön käyttöä,

jotta ne olisivat saavutettavissa mahdollisimman suurelle joukolle työnhakijoita. Tämä havainto korostaa selitettävyyden tärkeyttä osana tekoälypohjaisten järjestelmien läpinäkyvyyttä ja käyttäjäystävällisyyttä. Yksi haastateltavista kuvasi seuraavasti, miten tekoälypohjaisten rekrytointijärjestelmien päätöksiä ja toimintoja tulisi selittää:

”En usko, että työnhakijan pitäisi olla tekninen asiantuntija ymmärtääkseen tekoälyn päätöksiä. Selitysten tulisi olla kirjoitettu selkokielellä, ilman monimutkaisia teknisiä yksityiskohtia.” (H5)

Tulosperusteisuus nousi keskeisinä selitystyyppinä työnhakijoiden odotuksissa tekoälypohjaisten rekrytointijärjestelmien selitettävyydestä. Haastatellut painottivat, että tulosperusteisuus selitystyyppinä on erityisen tärkeää, sillä siinä korostuvat päätöksiin johtaneiden tekijöiden perusteltavuus ja tietopohjaisuus. Tämän tyyppiset selitykset auttavat työnhakijoita ymmärtämään selkeästi, miksi heidät valittiin tai miksi he eivät edenneet rekrytointiprosessissa. Tulosperusteiset selitykset nähdään konkreettisenä palautteena, mikä auttaa työnhakijaa tunnistamaan hakemuksensa vahvuuksia ja heikkouksia, ja siten hyödyntämään tätä informaatiota tulevia hakemuksia varten. Yksi haastateltavista kuvasi tulosperusteista selitettävyyttä seuraavasti:

”Henkilökohtaisesti koen molemmat tärkeiksi eri syistä. Tulosperusteinen selitys antaa minulle selkeämmän vastauksen siihen, miksi valittiin tai hylättiin. Kun taas prosessiperusteisten selitykset auttaisivat ymmärtämään koko rekrytointiprosessin kokonaisuutta tarkemmin. Hakemukseen liittyvät tulokset ovat kuitenkin hyödyllisempiä, koska niistä saa enemmän irti” (H6)

Prosessiperusteiset selitykset nähdään myös olennaisina päätöksentekoprosessin ymmärtämisen kannalta, vaikka tulosperusteisuus nähdään selkeästi hyödyllisempänä näkökantana. Selitystyyppinä prosessiperusteiset selitystyyppit kohdistuvat oikeudenmukaisuuden, turvallisuuden, suorituskyvyn ja tekoälyn vastuullisuuden näkökulmiin. Monet haastatelluista toivoivat selityksiä, kuinka varmistetaan, ettei

tekoälyjärjestelmä syrji tai tee virheellisiä päätöksiä. Oikeudenmukaisuuden varmistamiseksi ehdotettiin tilasto- ja vertailutietojen esittämistä, ulkopuolisen tahon validointia tekoälyjärjestelmän toimivuudesta ja konkreettisten toimenpiteiden esittämistä syrjinnän estämiseksi. Yksi haastateltavista kuvasi prosessiperusteista selitettävyyttä seuraavasti:

”Tämä voisi juuri näkyä, vaikka niissä vertailutiedoissa. Esimerkiksi siinä voisi olla esitettyä millaista sukupuolijakaumaa on ollut. Olisi muuten mielenkiintoista nähdä, kuinka paljon miehiä ja naisia on valittu eteenpäin tuleviin vaiheisiin.” (H1)

5.1.1 Tuloperusteisen selitettävyyden määrittäminen

Henkilökohtaisten tarpeiden ja hakemuksen ominaisuuksien huomioiminen tekoälypohjaisissa rekrytointijärjestelmissä oli toinen keskeinen havainto aineistosta. Työnhakijat toivovat, että tekoälyjärjestelmien kautta huomioidaan heidän yksilöllisiä tarpeitansa palautteen muodossa, joka sisältää hakemuksen sisältöön liittyviä tekijöitä. Haastateltavat arvostavat, että rekrytinnissa käytetty tekoälyjärjestelmä pystyisi analysoimaan hakemuksia syvällisemmin ja tarjoamaan yksilöllistä palautetta hakemuksen vahvuuksista ja mahdollisista kehityskohteista. Tämä auttaa työnhakijoita ymmärtämään, miten hänen hakemuksensa on suoriutunut. Yksi haastateltavista kuvasi hakemuksien ominaisuuksien huomioimisesta seuraavasti:

”Olisihan se mukavaa nähdä selkeästi jäsennellyn yhteenvedon siitä, miten pisteytykset muodostuivat eri kriteerien perusteella.” (H1)

Tuloperusteiset selitykset koetaan erityisen tärkeäksi tilanteissa, joissa hakija on päässyt pitkälle rekrytointiprosessissa, mutta ei tullut valituksi. Näissä tilanteissa työnhakijat toivovat yksityiskohtaisempaa palautetta siitä, missä he onnistuivat ja missä jäivät jälkeen. Selityksissä tulee olla tietopohjaista perusteltavuutta, joka auttaa hakijaa ymmärtämään päätöksen takana olevat kriteerit ja mahdollisuudet puutteet, jotka

vaikuttivat lopulliseen päätökseen. Yksi haastateltavista kuvasi perusteltavuuden tärkeyttä seuraavasti:

”Haluaisin tietää, miksi tietty kokemus tai taito oli ratkaiseva. Jos saan tietoa siitä, mitä puutteita hakemuksessani oli, voin kehittää itseäni seuraavaa hakemusta varten.” (H5)

Tulosperusteiset selitykset nähdään myös keinona tehdä rekrytointiprosessista läpinäkyvämpää ja oikeudenmukaisempaa. Selkeät perustelut auttavat työnhakijoita hyväksymään rekrytointiprosessin tulokset ja mahdollisesti voivat vähentää epävarmuutta tai turhautumista, joka saattaa syntyä erityisesti tilanteissa, jossa työnhakija jää ilman selkeää palautetta. Ymmärrys siitä, millä perusteilla päätökset tehdään, lisää hakijoiden luottamusta sekä rekrytointijärjestelmiin että työnantajiin. Yksi haastateltavista kuvasi tulosperusteista selitettävyyttä seuraavasti:

”Jos saan tietoa siitä, miksi hakemukseni ei päässyt jatkoon. Tiedän tarkemmin, mihin minun pitää keskittyä seuraavalla kerralla. Se auttaa ymmärtämään, mitä juuri tässä prosessissa painotettiin.” (H3)

Selitysten hyödyntäminen rekrytointiprosessissa avaa myös oppimisenäkökulman, jonka kautta työnhakijat voivat kehittää osaamistaan ja hakutaitojaan tulevaisuutta varten. Haastateltavat korostivat, että konkreettinen palaute siitä, missä heidän hakemuksensa suoriutui hyvin ja missä jäi vajaaksi, tarjoaa mahdollisuuden oppia ja kasvaa työnhakijana. Tämä oppimisprosessi nähdään erityisen tärkeänä, kun hakemuksissa toistuvat tietyt puutteet tai kun hakijat haluavat parantaa osaamistaan tiettyjen taitojen tai kokemusten osalta.

Tulosperusteiset selitykset, jotka sisältävät oppimisenäkökulman kannalta hyödyllistä tietoa, voivat auttaa työnhakijoita ymmärtämään, mihin heidän kannattaa keskittyä ammatillisessa kehityksessään. Tämä auttaa työnhakijoita tunnistamaan paremmin

millaisia taitoja ja kokemuksia työnantajat arvostavat eniten, mikä johtaa heidät lisäämään valmiuksiaan ja ottamaan huomioon omia erityispiirteitä tarkemmin tulevia hakemuksia varten. Yksi vastaajista kuvasi oppisnäkökulmaa seuraavasti:

”Se olisi hyödyllistä, jos saisin tietää, mitkä osa-alueet hakemuksestani olivat vahvoja ja mitkä kaipaavat kehitystä. Tämä auttaisi kehittymään ja parantamaan tulevia hakemuksia.” (H2)

Vertailutiedot nousivat merkittäväksi osa-alueeksi työnhakijoiden odotuksissa tekoälypohjaisten rekrytointijärjestelmien selitettävyydessä. Haastateltavat korostivat, että vertailutiedot muihin hakijoiden nähden tarjoaisivat heille arvokasta tietoa omasta suoriutumisestaan ja kilpailukyvyystään. Monille työnhakijoille on tärkeää ymmärtää, miten heidän hakemuksensa vertautuu muiden hakijoiden osaamiseen ja kokemukseen, jotta he voivat arvioida omaa asemaansa työnhakumarkkinoilla. Tämä vertailu koettiin reiluna ja motivoivana, sillä se auttaa tunnistamaan omat vahvuudet ja kehityskohteet.

Vertailutiedot voivat käsittää erilaisia hakemusten arvioinnissa käytettyjä mittareita, kuten hakijoiden koulutustasoa, työkokemuksen määrää tai tiettyjä taitoja, joita tekoälyjärjestelmä painottaa. Haastateltavat ilmaisivat, että tällaiset tiedot voivat tarjota heille konkreettista suuntaa siihen, miten heidän kannattaa jatkossa parantaa omaa osaamistaan ja tuoda esille vahvuuksiaan hakemuksissaan. Yksi haastateltavista kuvasi kiinnostustaan vertailutietojen yksityiskohtiin seuraavasti:

”Olisi kiinnostavaa tietää, miten pärjäsini muihin hakijoihin nähden, esimerkiksi millainen koulutustaso heillä oli tai kuinka paljon kokemusta heillä oli verrattuna omaani. Tämä auttaisi ymmärtämään, missä voin parantaa.” (H3)

Tulosperusteinen selitettävyys korostuu työnhakijoiden näkemyksissä tekoälypohjaisten rekrytointijärjestelmien käytössä. He toivovat selkeää ja konkreettista palautetta siitä, miksi valittiin tai miksi he eivät edenneet rekrytointiprosessissa. Tämä auttaa

ymmärtämään päätösten takana olevat tekijät sekä tunnistamaan vahvuuksia ja heikkouksia. Tulospohjaiset selitykset lisäävät myös luottamusta ja läpinäkyvyyttä rekrytointiprosessiin, mikä vähentää epävarmuutta työnhakijoissa. Oppimiskäytäntö on myös keskeinen, sillä selitykset tarjoavat työnhakijoille mahdollisuuden kehittää omaa osaamistaan ja hakutaitojaan tulevaisuutta varten. Vertailutiedot muihin hakijoihin nähden auttavat työnhakijoita ymmärtämään omaa asemaansa työmarkkinoilla ja tunnistamaan kehityskohteensa kilpailukykyä vahvistamiseksi.

5.1.2 Prosessiperusteisen selitettävyyden määrittäminen

Prosessiperusteinen selitettävyys nousi esiin tutkimuksen tuloksissa erityisesti silloin kun, työnhakijat pohtivat tekoälypohjaisten rekrytointijärjestelmien päätöksenteon vastuullisuutta, oikeudenmukaisuutta ja turvallisuutta. Vaikka tulospohjaiset selitykset koettiin konkreettisiksi ja hyödyllisiksi, monet haastatelluista toivat esille myös tarpeen ymmärtää, miten päätökset on tehty ja mitkä prosessit ovat olleet osallisena päätöksenteossa. Tämä prosessin avaaminen nähtiin tärkeänä erityisesti silloin, kun työnhakija haluaa varmistua siitä, että tekoälytoimintaa ohjaavat reilut ja läpinäkyvät periaatteet.

Vastuullisuus nousi merkittäväksi näkökulmaksi, sillä työnhakijat haluavat tietää, millaisia periaatteita ja toimintamalleja tekoäly noudattaa. Monet haastateltavat korostivat, että tekoälyjärjestelmän käytössä on tärkeää varmistaa, ettei se syrji ketään tiettyjen ominaisuuksien, kuten iän, sukupuolen tai taustan perusteella. Tämä vaatii järjestelmiltä konkreettisia mekanismeja ja selityksiä siitä, miten ennakoasenteita tunnistetaan ja vältetään. Yksi haastateltavista kuvasi vastuullisnäkemyksiä seuraavasti:

”Olisi hienoa tietää, että tekoäly toimii oikeudenmukaisesti, eli että se ei syrji ketään perusteettomasti. Jonkinlainen vahvistus siitä, että kaikki hakijat arvioidaan samalla tavalla, toisi luottamusta järjestelmään.” (H6)

Oikeudenmukaisuus liittyi myös vertailutietojen esittämiseen, sillä työnhakijat halusivat varmistua, että tekoäly tekee päätökset tasapuolisesti ilman ennakkoluuloja. Erityisesti he toivoivat, että järjestelmä avaisi päätöksentekoprosessia tavalla, joka antaa heille mahdollisuuden arvioida omaa menestystään suhteessa muihin. Yksi haastateltavista kuvasi vertailutietojen esittämistä seuraavasti:

”Olisi mielenkiintoista nähdä tilastot siitä, miten hakijoiden erilaisia taustoja ja osaamista on arvioitu, jotta varmistutaan, ettei tekoäly syrji ketään.” (H1)

Vaikutukset sekä yksilöön että yhteiskuntaan saivat myös vähän huomiota haastateltavilta. Osa työnhakijoista haluaa tietää, miten tekoälyn tekemät vaikuttavat heidän työllistymismahdollisuuksiinsa ja miten nämä päätökset voivat heijastua laajemmin työmarkkinoihin ja rekrytointikäytäntöihin. Osa haastateltavista odottaa, että tekoälyjärjestelmä avaisi myös näitä laajempia vaikutuksia päätöksenteon taustalla. Yksi haastateltavista kuvasi mielenkiintoaan tekoälyn vaikutuksista seuraavasti.

”Olisi mielenkiintoista tietää, miten tekoäly vaikuttaa työnhakijoiden asemaan laajemmin. Mitä vaikutuksia sillä on työmarkkinoihin ja miten sen käyttöönotto voi muuttaa rekrytointia tulevaisuudessa.” (H3)

Yhteenvetona voidaan todeta, että prosessiperusteinen selitettävyys kattaa monia työnhakijoille tärkeitä ulottuvuuksia, jotka liittyvät tekoälyn toiminnan avoimuuteen, vastuullisuuteen ja turvallisuuteen. Selkeät ja kattavat selitykset tekoälyn toimintaperiaatteista edistävät työnhakijoiden luottamusta järjestelmään ja parantavat heidän kokemustaan rekrytointiprosessista.

5.2 Tekoälyjärjestelmien vuorovaikutuksellisuuden merkitys rekrytointiprosessissa

Rekrytointiprosessiin liittyvän vuorovaikutuksellisen läpinäkyvyyden tärkeys korostui selvästi tutkimuksen tuloksissa. Tekoälyjärjestelmiä ei nähdä hakijoiden näkökulmasta

ainoastaan työkaluina hakemusten käsittelyyn ja analysointiin, vaan niiden odotetaan toimivan aktiivisessa vuorovaikutuksessa työnhakijoiden kanssa. Haastateltavat toivovat saavansa ajantasaista ja reaaliaikaista palautetta hakemustensa tilanteesta, ja heille on tärkeää olla tietoisia siitä, missä vaiheessa prosessi etenee. Tämä voi auttaa vähentämään hakuprosessin aikana koettua epävarmuutta ja parantaa rekrytointiprosessin läpinäkyvyyttä. Yksi haastateltavista totesi seuraavaa:

”Tärkeää olisi saada mahdollisimman nopeasti tietoa siitä, että hakemus on kunnossa tai missä vaiheessa prosessi menee. En tarvitse mitään syvällistä, mutta jokin yhteenveto tai lyhyt viesti olisi hyvä.” (H6)

Vuorovaikutus ilmenee pääsääntöisesti tiedon välittämisenä. Reaaliaikaisen tiedon saaminen siitä, että hakemus on vastaanotettu tai että se etenee seulonta vaiheesta seuraaviin vaiheisiin, koetaan tärkeäksi vuorovaikutuksen elementiksi. Lisäksi monille hakijoille on erityisen tärkeää, että tekoälypohjainen rekrytointijärjestelmä tarjoaa konkreettisia selityksiä hylkäysten taustalla olevista syistä, jotta he voivat oppia ja kehittää tulevia hakemuksiaan. Tämä vahvistaa hakijoiden sitoutumista ja parantaa käsitystään rekrytointiprosessista.

”Minusta olisi hyvä tietää, missä kohtaa prosessia tekoäly tekee päätöksiä ja milloin siirrytään inhimilliseen arviointiin. Se lisäisi luottamusta siihen, että hakemukseni käsitellään kunnolla.” (H4)

Vuorovaikutuksellisuus ei kuitenkaan rajoitu vain tiedon jakamiseen; se kattaa myös mahdollisuuden aktiiviseen osallistumiseen ja kysymysten esittämiseen. Osa hakijoista korostaa, että he haluaisivat esittää kysymyksiä tekoälylle prosessin aikana ja saada selvennyksiä, miten tietyt päätökset on tehty tai mitkä tekijät painottuvat hakemusten arvioinnissa. Samalla kuitenkin osa hakijoista pitää automaattista tiedottamista riittävänä ja vähemmän kuormittavana, mikä viittaa siihen, että tekoälypohjaiselta rekrytointijärjestelmältä odotetaan mukautuvuutta erilaisten tarpeiden mukaan.

”Olisi hyvä, jos tekoäly selittäisi automaattisesti, mitä päätöksen taustalla on, ettei erikseen tarvitse kysyä joka vaiheessa lisätietoja.” (H3)

Selkeys tekoälyn ja ihmisen roolien välillä on myös merkittävä osa vuorovaikutteisuutta rekrytointiprosessissa. Monille hakijoille on tärkeää tietää, milloin tekoäly tekee päätöksiä ja milloin inhimillinen arviointi astuu mukaan. Tämä tuo prosessiin luotettavuutta ja läpinäkyvyyttä, mikä auttaa hakijoita ymmärtämään paremmin koko rekrytointiprosessiin logiikkaa ja päätöksenteon taustalla vaikuttavia tekijöitä.

5.2.1 Reaaliaikaisen informaation merkitys

Tutkimuksen tuloksissa nousee vahvasti esille työnhakijoiden tarve saada reaaliaikaista ja välitöntä palautetta hakemusten tilasta tekoälypohjaisissa rekrytointijärjestelmissä. He toivovat, että tekoälyjärjestelmä tarjoaisi selkeää ja ajankohtaista tietoa siitä, missä vaiheessa hakemus on rekrytointiprosessin aikana. Monille oli tärkeää, että he voisivat seurata hakemuksensa etenemistä ja saada tarvittaessa nopeasti tiedon siitä onko hakemuksessa puutteita tai miten se on suoriutunut kriteerien valossa. Yksi haastateltavista kuvasi tarvetta reaaliaikaiselle tiedottamiselle seuraavasti:

”Haluan säännöllistä ja selkeää kommunikointia alusta alkaen, esimerkiksi vahvistuksen siitä, että CV on vastaanotettu. Toivon sanallisempaa viestintää heti hakemuksen jättämisestä lähtien ja haluan tietoa rekrytointiprosessin vaiheista.”
(H2)

Hakijat toivovat erityisesti, että he saisivat nopeasti palautetta siitä, missä kohtaan hakemuksen käsittelyssä ollaan, ja pystyisivät näin seuraamaan rekrytointiprosessin eri vaiheita. Tämä voisi olla esimerkiksi vahvistus hakemuksen vastaanottamisesta, tieto hakemuksen siirtymisestä seulonnan läpi tai milloin hakemuksen arviointi siirtyy ihmisen käsiteltäväksi. Erityisen tärkeäksi koettiin mahdollisuus saada reaaliaikaista palautetta siinä vaiheessa, jos hakemus hylätään. Monissa tapauksissa hakijat kaipaivat tietoa siitä,

miksi heidän hakemuksensa ei päässyt jatkoon, jotta he voisivat oppia ja kehittyä seuraavia hakuja varten.

Yksi haastateltava mainitsi, että reaaliaikainen vuorovaikutus tekoälyjärjestelmän kanssa voisi toteutua esimerkiksi chat-ominaisuuden kautta, jolloin hakija voisi saada vastauksia kysymyksiin nopeasti ja joustavasti. Tämä toive vuorovaikutteisuudesta korostaa, kuinka tekoälyn käyttöä ei nähdä vain teknisenä analysoinnin työkaluna, vaan myös dialogisena vuorovaikutustyyppinä, jossa rekrytointiprosessin kommunikaatio voidaan nopeuttaa ja mukauttaa käyttäjien tarpeiden mukaan. Tekoälyn nähdään mahdollisesti tarjoavan hakijoille nopeampaa, selkeämpää ja mukautuvampaa tiedottamista rekrytointiprosessiin. Yksi haastateltavista kuvasi kommunikaatiota seuraavasti:

”Toivoisin oikeaa aikaista kommunikaatiota, niin se säästää omia resursseja. Tekoäly voisi jopa toimia nopeammalla tavalla kommunikoinnin suhteen kuin ilman tekoälyä. Se voisi olla semmoinen chatti tekoälyn kanssa.” (H2)

Tämä sitaatti tiivistää työnhakijoiden odotukset nopeasta ja helposti saatavilla olevasta viestinnästä tekoälypohjaisten rekrytointijärjestelmien suhteen. Reaaliaikainen kommunikointi nähdään tärkeänä osana tekoälypohjaisten rekrytointijärjestelmien tehokasta toimintaa ja hakijoiden odotuksia tukevana elementtinä. Se tuo prosessiin paitsi nopeutta, myös joustavuutta, jolloin hakijat voivat tuntea olonsa varmemmaksi ja tietoisemmaksi koko hakuprosessin ajan.

5.2.2 Mahdollisuus kysyä lisätietoja ja tarkennuksia

Mahdollisuus kysyä tarkennuksia tekoälypohjaisen rekrytointijärjestelmän tekemistä päätöksistä nousi yhdeksi keskeiseksi toiveeksi tutkimuksen tuloksista työnhakijoiden keskuudessa. Useat työnhakijat eivät halua jäädä passiivisiksi prosessien seuraajiksi, vaan he toivovat mahdollisuutta aktiiviseen vuorovaikutukseen, jolloin he voivat ymmärtää paremmin hakuprosessin aikana tehtyjä päätöksiä. Toisaalta osa

haastateltavista näkee automaattisen tiedottamisen helpompana ja vähemmän kuormittavana vaihtoehtona lisätietojen sekä tarkennuksien esittämiseen.

Tarve kysyä lisätietoja ja tarkennuksia korostui erityisesti negatiivisten rekryointipäätöksien yhteydessä. Haastateltavat kertoivat usein siitä, kuinka hylkäysilmoitukset jäivät tyhjänpäiväisiksi, ilman konkreettista selitystä siitä, miksi juuri heidän hakemuksensa ei täyttänyt vaatimuksia tai mitä he voisivat kehittää tulevaisuudessa. Monissa kommentteissa korostettiin, että pelkkä tieto hylkäyksestä ei riitä – työnhakijat haluavat tietää syyt päätösten taustalla, sillä he käyttävät itsekin paljon aikaa työhakemuksien tekemiseen.

”Minulle olisi tärkeää tekoälyn suhteen, että saisi sanallisen palautteen tai tietoa siitä, mikä vaikutti mihinkin päätöksessä. Kaikista olennaisinta olisi saada se palaute, koska ilma sitä jää tyhjän päälle tai jää epäonnistunut fiilis.” (H2)

Aineistossa ilmeni myös, että osa suosii automaattista tiedottamista, ilman että heidän tarvitsisi itse aktiivisesti kysyä lisätietoja tai tehdä ylimääräisiä ponnisteluja saadakseen selityksiä tekoälypohjaisten rekryointijärjestelmien päätöksistä. Tämä näkemys nousi esiin erityisesti niissä kommentteissa, jossa hakijat halusivat saada selkeää, mutta samalla yksinkertaista ja helposti ymmärrettävää tietoa automaattisesti päätöksenteon eri vaiheista.

Osa haastateltavista kokivat, että tällainen automaattinen tiedottaminen tekisi prosessista sujuvamman ja vähentäisi heidän omaa taakkaansa. Heille oli tärkeää, että rekryointipohjainen tekoälyjärjestelmä tarjoaisi relevantteja tietoja automaattisesti, jolloin työnhakijan ei tarvitsisi erikseen ryhtyä selvittämään päätösten taustoja tai vaatia lisätietoja. Tämä lähestymistapa nähtiin erityisen hyödyllisenä silloin, kun hakuprosessi on jo ennestään kuormittava tai kun hakija joutuu käymään läpi useita hakemuksia ja rekryointiprosesseja samanaikaisesti. Tällöin automaattinen tiedottaminen toimi myös

stressin ja turhautumisen vähentäjänä. Yksi haastateltavista kuvasi automaattista tiedottamista seuraavasti:

”Kyllä, mutta vielä parempi olisi, jos tarvittavat lisäselvitykset ja tarkennukset annettaisiin automaattisesti, ilman että minun täytyisi erikseen niitä pyytää.” (H4)

Nämä näkökulmat aineistoista osoittavat, että osa työnhakijoista haluaa aktiivista vuorovaikutusta ja mahdollisuutta kysyä lisätietoja, kun taas toiset näkevät automatisoidun tiedottamisen riittävänä ja toivottavana ratkaisuna. Näin ollen molempien näkökulmien huomioiminen – sekä reaaliaikaisen palautteen ja tarkennusten mahdollistaminen että automaattinen tiedottaminen – ovat tärkeitä, jotta tekoälypohjainen rekryointijärjestelmä voi vastata moninaisten työnhakijoiden odotuksiin ja tarpeisiin mahdollisimman hyvin.

5.2.3 Tekoälyn ja ihmisen roolien selkeys

Tekoälyn ja ihmisen roolien selkeys rekryointiprosessissa nousi aineistossa esiin yhtenä työnhakijoiden keskeisenä toiveena. Hakijat kaipasivat tarkempaa ymmärrystä siitä, missä vaiheissa tekoäly tekee päätöksiä ja missä vaiheessa ihmisen rooli astuu mukaan prosessiin. Tämä selkeys tekoälyn ja ihmisten roolien välittämisessä nähtiin tärkeänä erityisesti siksi, että se auttaisi hakijoita hahmottamaan prosessin kokonaisuutta ja tarjoaisi heille mahdollisuuden arvioida, missä kohtaa he voivat odottaa inhimillistä arviointia ja milloin päätökset perustuvat algoritmien toimintaan.

Monet haastateltavat kokivat, että tekoälyn ja ihmisten roolien selkeyttäminen lisäisi prosessin läpinäkyvyyttä ja auttaisi heitä ymmärtämään paremmin rekryointiprosessin logiikkaa. Selkeä roolijako ja tiedottaminen siitä, milloin tekoäly tekee päätöksiä ja milloin ihminen osallistuu rekryointiprosessiin, lisäisi myös hakijoiden luottamusta siihen, että päätökset ovat oikeudenmukaisia ja reiluja. Yksi haastateltavista totesi seuraavasti:

”Haluaisin tietää, missä kohtaan tekoäly käsittelee hakemustani ja milloin päätöksenteko siirtyy ihmiselle. Se auttaisi minua ymmärtämään, missä vaiheessa voin odottaa inhimillistä arviointia ja kuinka paljon tekoäly on vastuussa valinoista.”
(H5)

Tekoälypohjaisten rekrytointijärjestelmien käyttö rekrytoinnissa tuo mukanaan tiettyjä ongelmia, jotka liittyvät erityisesti prosessin läpinäkyvyyteen. Hakijat korostivat, että ilman selkeää tietoa siitä, kuka tekee päätökset ja millä perusteilla, prosessi voi tuntua mekaaniselta ja epäinhimilliseltä. Erityisesti tilanteissa, joissa ei päästä jatkoon rekrytointiprosessissa, ilmaistiin toiveita saada tietää, onko hakijoiden hakemuksia käsitellyt tekoäly vai ihminen. Tämä lisäisi luottamusta rekrytointiprosessiin ja antaisi paremman käsityksen siitä, että kuinka paljon painoarvoa tekoälyn päätöksille annetaan ja milloin inhimillinen harkinta tulee mukaan.

Hakijat kaipasivat myös varmuutta siitä, että päätöksenteko on tasapuolista ja oikeudenmukaista riippumatta siitä, tehdäänkö päätökset tekoälyn vai ihmisen toimesta. Tämä voidaan saavuttaa vain, jos hakijoille kerrotaan selkeästi, missä vaiheessa tekoäly osallistuu prosessiin ja milloin ihmisen asiantuntemus tulee mukaan. Haastateltavat näkivät tämän tärkeänä osana koko rekrytointiprosessin läpinäkyvyyttä:

”Kyllä, olisi hyvä tietää, milloin tekoäly on vastuussa ja milloin päätös tulee ihmiseltä. Tämä auttaisi hahmottamaan, miten prosessi etenee ja milloin odottaa inhimillistä arviointia.” (H6)

6 Toteutus

Tässä tutkimuksessa konstruktion toteutuksessa aloitetaan tutkimuksessa esiin nousseista tarpeista, jotka liittyvät tekoälypohjaisten rekryointijärjestelmien käyttöön ja työnhakijoiden näkemyksiin niiden selitettävyydestä ja vuorovaikutteisuudesta. Empiirisen aineiston perusteella muodostetaan ratkaisu, joka vastaa pyrkii vastaamaan näihin odotuksiin ja kehittää rekryointiprosessia läpinäkyvämmäksi, oikeudenmukaisemmaksi ja käyttäjäystävällisemmäksi. Toteutusprosessi muodostuu seuraavista vaiheista: ensin määritellään motivaatiot ja vaatimukset, minkä jälkeen siirrytään suunnitteluvaiheeseen ja lopulta ratkaisun toteuttamiseen sekä arviointiin.

6.1 Motivaatio ja vaatimukset

Motivaatio konstruktion toteuttamiselle pohjautuu tutkimuksen keskeisistä tuloksista, joissa työnhakijoiden odotuksia ja tarpeita, jotka liittyvät tekoälypohjaisten rekryointijärjestelmien läpinäkyvyyteen ja vuorovaikutuksellisuuteen. Työnhakijat kokevat erityisesti päätösten selitettävyyden ja saamisen ongelmakohtaksi, minkä vuoksi tekoälypohjaisia rekryointijärjestelmiä tuli kehittää selkeämmäksi ja käyttäjäystävällisemmiksi hakijoita ajatellen. Alla on kuvattu tarkemmin kaksi pääteemaa ja niihin liittyvät vaatimukset.

Tutkimuksen tuloksissa korostuu työnhakijoiden tarve ymmärtää, miten tekoälypohjaiset rekryointijärjestelmät tekevät päätöksiä rekryointiprosessin aikana. Selitettävyyden odotukset liittyvät ennen kaikkea päätöksentekoprosessiin selkeyteen, selitysten yksinkertaisuuteen ja konkreettisiin perusteluihin, miksi tietty päätös on tehty. Työnhakijat odottavat, että tekoälyjärjestelmät tarjoaisivat selityksiä ilman monimutkaista termistöä. He haluavat tietää, miksi he eivät edenneet prosessissa tai miksi joku toinen valittiin heidän sijastaan. Haastateltavat toivoivat konkreettisia tulosperusteisia selityksiä, jotka selittävät mitä tekijöitä rekryoinnin arvioinnissa painotettiin ja miten heidän hakemuksensa pärjasi suhteessa näihin kriteereihin. Työnhakijat arvosta myös selityksiä, jotka avaavat päätöksentekoprosessia

kokonaisuutena. Tämä voisi kattaa esimerkiksi sen, miten tekoäly varmistaa oikeudenmukaisuuden ja syrjimättömyyden päätöksenteossa. Alla olevassa taulukossa 2 on luetteloituna vaatimukset selitettävyydelle.

| Vaatus | Kuvaus |
|--------|--------------------------------------------------------------------|
| 1 | Selkeät ja yksinkertaistetut selitykset |
| 2 | Tulosperusteisten selitysten tarjoaminen |
| 3 | Prosessiperusteisten selitysten tarve kokonaisuuden ymmärtämiseksi |

Taulukko 2. Vaatimukset selitettävyydelle.

Vuorovaikutus työnhakijan ja tekoälypohjaisen rekrytointijärjestelmän välillä on toinen keskeinen teema tutkimuksessa. Työnhakijat haluavat aktiivista ja joustavaa vuorovaikutusta tekoälyjärjestelmän kanssa, ja heille on tärkeää, että se ei ole ainoastaan päätöksiä tekevä työkalu, vaan myös reaaliaikaisena palautetta antavana kommunikoivana yhteistyökumppanina. Vuorovaikutuksen odotetaan olevan myös kaksisuuntaista, joten on tärkeää, että on mahdollisuus esittää kysymyksiä ja saada selityksiä tekoälyn tekemistä päätöksistä, erityisesti hylkäysten yhteydessä. Lisäksi tekoälypohjaisen rekrytointijärjestelmän kannalta nähdään tärkeänä ominaisuutena automaattinen tiedottaminen palautteiden suhteen, mutta huomioitavaa on myös, että automaation ja ihmisten roolit rekrytointiprosessissa ovat selkeitä työnhakijalle. Tämä tarkoittaa vuorovaikutuksessa, että työnhakijalle viestitään selkeästi, missä vaiheessa prosessia työhakemus on ja kuka käsittelyä hoitaa – tekoälyjärjestelmä vai ihminen – jolloin työnhakijalla säilyy ymmärrys automaation ja ihmisten rooleista rekrytointiprosessissa. Alla olevassa taulukossa 2 on esitettyä vuorovaikutukselliset vaatimukset.

| Vaatimukset | Kuvaus |
|-------------|-----------------------------------------------|
| 1 | Reaaliaikainen tiedottaminen |
| 2 | Mahdollisuus kysyä lisätietoja |
| 3 | Automatisoitu tiedottaminen |
| 4 | Automaatio ja ihmisten roolien selkeyttäminen |

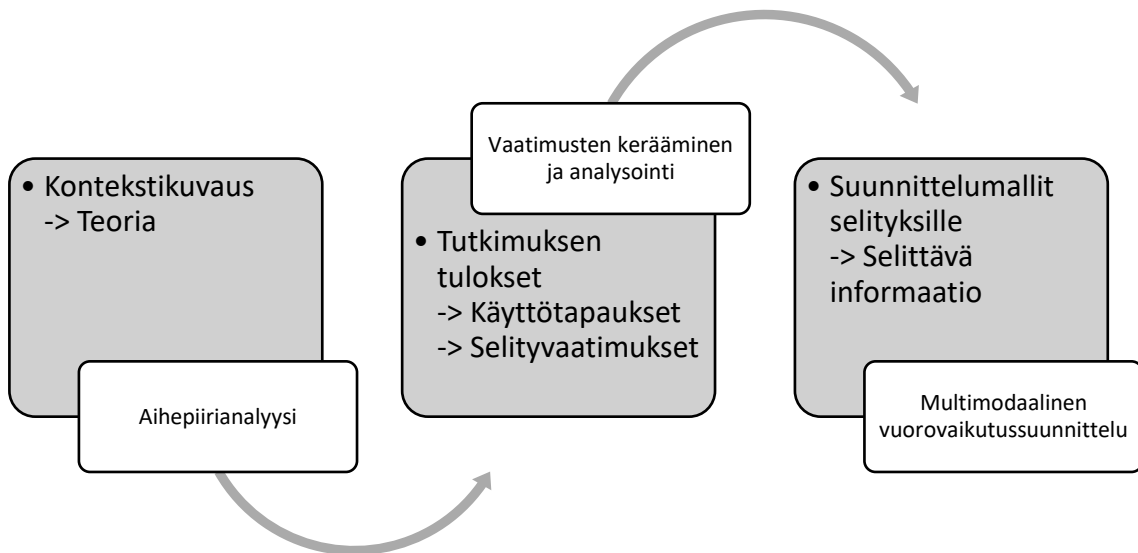
Taulukko 3. Vuorovaikutukselliset vaatimukset.

Tämän tutkielman konstruktion motivaationa on luoda luova ja heuristinen ohjeistus, joka antaa rekrytointijärjestelmien kehittäjille ja suunnittelijoille konkreettisia suuntaviivoja tekoälypohjaisten rekrytointijärjestelmien läpinäkyvyyden ja vuorovaikutuksellisuuden parantamiseen työnhakijoita ajatellen. Ohjeistuksen suunnittelussa huomioidaan erityisesti aineistosta nousseet työnhakijoiden tarpeet ja odotukset sekä hyödynnetään tutkimuksen teoriasta saatua tietopohjaa. Ohjeistuksen rakenne jakautuu kahteen pääosioon, jotka ovat selitettävyyden ja vuorovaikutteisuuden. Lisäksi ohjeistuksessa huomioidaan rekrytointiprosessien ja -järjestelmien elementtejä, kuten sen standardisoituja toimintatapoja sekä työnhakijoista käytettäviä tietotyyppisiä. Valitsin ohjeistuksen esittämistavaksi infografiikan, sillä se mahdollistaa tärkeimpien asioiden selkeästi ja tiiviisti rekrytointiprosessien- ja järjestelmien suunnittelijoille.

6.2 Suunnittelu

Tämän tutkielman ohjeistuksen suunnitteluprosessi seuraa osittain aikaisemmin teoriassa mainitun DoReMi-suunnittemallin käytäntöä. DoReMi-suunnittelumalli sisälsi kolme päävaihetta, joita ovat aihepiirianalyysi, vaatimusten kerääminen ja analysointi sekä multimodaalinen vuorovaikutussuunnittelu. Aihepiirianalyysi tarkoitti kontekstikuvauksen muodostamista. Vaatimusten kerääminen ja analysointi käyttötapausten sekä selitysvaatimusten määrittämistä. Multimodaalinen vuorovaikutussuunnittelu suunnittelumallien luomista selityksille.

Tässä tutkielmassa DoReMi-suunnittelumallin osalta huomioidaan aihepiirianalyysin kontekstikuvauksessa teoriasta kerättyä tietoa tutkimukseen liittyvästä aiheesta. Vaatimusten kerääminen ja selitysvaatimusten määrittäminen muodostui tutkimuksen empiirisen tutkimuksen tuloksista ja aineiston analyysistä, josta saatiin tietoa tutkimuksen teemojen ja tutkimuskysymysten mukaisesti selitettävyyteen ja vuorovaikutukseen liittyviä vaatimuksia sekä käyttötapauksia. Multimodaalisessa vuorovaikutussuunnittelussa ei huomioida XAI-menetelmien sisältämää teknistä näkökulmaa, sillä selitysmallien tekninen konseptuaalinen suunnittelu ei ole keskeinen osa-alue tutkimuksessa. Multimodaalisessa vuorovaikutussuunnittelussa keskitytään luomaan konseptuaalisia ja prototyyppisiä selityksiä, jotka sisältävät ominaisuuksia tai toiminnallisuuksia tutkimuksen tuloksista nousseitten teemojen mukaisesti. Tämä tarkoittaa tarkemmin kontekstinmukaista ja selitysvaatimukset täyttävää selittävän informaation suunnittelua haastateltaville. Alla oleva kuvio X esittää DoReMi-suunnittelumallin käyttöä tämän tutkimuksen suunnittelussa.



Nimi: DoReMi-suunnittelumalli tässä tutkimuksessa.

Seuraavaksi suunnitteluprosessissa tarkastellaan yhteyksiä teoriaan liittyvän kontekstikuvauksen ja tutkimuksen tulosten välillä. Tämä tarkoittaa, että yhdistetään tekoälypohjaisiin rekryointijärjestelmiin liittyvä konteksti yksityiskohtaisemmin tutkimuksen tuloksista ilmenneisiin käsitteellisiin selitettävyyden ja vuorovaikutuksellisuuden odotuksiin hakijoiden kannalta. Tuloksista saadun tiedon perusteella keskeisiksi selitettävyysodotuksiksi nousivat konkreettisten valintakriteerien avaaminen, tulosten vertailutiedot ja kehittävän palautteen antaminen. Vuorovaikutukselliseksi odotuksiksi muodostuivat reaaliaikainen ja automaattinen tiedottaminen, mahdollisuus kysyä lisätietoja sekä automaation ja ihmisten roolien selkeyttäminen. Tuloksissa selviää myös, että keskeisimmät selitettävyysodotukset vaikuttavat myös haastateltavien prosessiperusteiseen läpinäkyvyyden tulkintaan tekoälypohjaisesta rekryointijärjestelmästä kokonaisuudessa, esimerkiksi oikeudenmukaisuuden ja vaikutuksien selitettävyyden näkökulmista.

Seuraavaksi suunnitteluprosessissa muodostetaan selitys- ja vuorovaikutusvaatimukset kuvauksineen edellisessä kappaleessa mainittuihin osa-alueisiin. Haastateltavien mukaan konkreettisten valintakriteerien avaaminen tarkoittaa, että tekoälypohjaisen rekryointijärjestelmän tulisi selkeästi selittää, mitkä tekijät vaikuttivat rekryointipäätökseen ja millä tavalla hakijan ominaisuudet vastasivat asetettuja kriteerejä. Tämä tarkoittaa, että esimerkiksi hakijan kokemus, koulutus tai osaamisalueiden tulisi olla ymmärrettävästi avattuna, jolloin työnhakija pystyy hahmottamaan, missä määrin ne vastaavat vaatimuksia.

Tulosten vertailutiedot tarkoittaa, että tekoälypohjaisen rekryointijärjestelmän tulee tarjota vertailutietoja työnhakijan sijoittumisesta suhteessa muihin hakijoihin eri kriteereissä. Haastateltavien mukaan vertailutietojen tarjoaminen antaa mahdollisuuden ymmärtää omaa sijoittumistaan kokonaisuudessaan. Tämä selkeyttää hakijoiden kannalta pisteytymallin toimintaa ja lisää siten myös läpinäkyvyyttä prosessiperusteisesti.

Kehittävä palaute kehityksen tueksi tarkoittaa, että tekoälypohjaisen rekrytointijärjestelmän tulisi tarjota palautetta työnhakijan osa-alueista, joissa hän voisi kehittyä tulevaisuutta ajatellen. Haasteltavien näkökulmasta palaute voisi tukea ammatillisen kehityksen suunnittelussa ja auttaa häntä parantamaan mahdollisuuksiaan tulevissa rekrytointiprosesseissa. Lisäksi palaute voi vahvistaa kokemusta rekrytointiprosessista, sillä se antaa työnhakijalle arvokasta tietoa omasta osaamisestaan ja osoittaa, että hänen potentiaalinsa ja kehityksen otetaan vakavasti.

Reaaliaikainen ja automaattinen tiedottaminen tarkoittaa, että tekoälypohjaisen rekrytointijärjestelmän tulisi viestiä hakijoille heidän hakemuksensa tilasta ja etenemisestä nopeasta ja joustavasti. Haastateltavien mukaan tämä voisi tarkoittaa esimerkiksi ilmoitusta hakemuksen vastaanottamisesta, päivityksiä seulontavaiheista sekä tietoa siitä, milloin hakemus siirtyy ihmisen arvioitavaksi. Hakijat toivoivat erityisesti tiedonsaantia siinä vaiheessa, jos hakemus ei etene, jolloin he voivat saada siihen johtavaa tietoa.

Mahdollisuus kysyä lisätietoja tekoälypohjaisten rekrytointijärjestelmän tekemistä päätöksistä nousi työnhakijoiden toiveissa esille. Osa työnhakijoista haluavat ymmärtää tarkemmin, miksi tietyt päätökset on tehty ja millä perusteilla he eivät edenneet prosessissa. Erityisesti hakijan kannalta negatiiviset päätökset, kuten hakemuksen hylkäys, nähtiin tilanteina, joissa lisätiedon saaminen koettiin olennaisiksi. Vaikka reaaliaikainen ja automaattinen tiedottaminen on haastateltavan näkökulmasta helpompaa, mahdollisuus kysyä lisätietoja nähdään silti yhtenä potentiaalisena vaihtoehtona, miten vuorovaikuttaa tekoälypohjaisen rekrytointijärjestelmän kanssa.

Automaatio ja ihmisten roolien selkeyttäminen rekrytointiprosessissa oli työnhakijoiden keskeinen toive. Hakijat halusivat tietää tarkemmin, missä vaiheissa tekoäly käsittelee hakemuksia ja milloin ihmisen arviointi tulee mukaan prosessiin. Tämän tiedon avulla he kokivat voivansa paremmin ymmärtää rekrytointiprosessin kokonaisuutta ja arvioida, milloin he voivat odottaa inhimillistä harkintaa. Selkeä roolijako myös lisäisi hakijoiden

luottamusta prosessiin, sillä se tarjoaisi näkyvyyttä siihen, että päätökset ovat oikeudenmukaisia ja perustuvat tarkoituksenmukaisiin arviointikriteereihin.

Seuraavaksi suunnitteluprosessissa yhdistetään edellä mainitut vaatimukset käyttötapauksiin, eli rekrytoinnin kannalta konseptuaalisiin tilanteisiin. Konkreettisten valintakriteerien avaaminen yhdistetään teorian tulkitsemisen perusteella syntyneeseen vaihtoehtoon, joka on tekoälypohjaisen rekrytointijärjestelmän ansioluettelon automaattinen analysointi. Tämä tarkoittaa hakemuksen jättämisen jälkeistä vaihetta, jossa hakijan tiedoista muodostetaan tuloksia kriteerien perusteella. Tässä on mahdollista välittää työnhakijalle mitkä tekijät vaikuttavat päätökseen ja millä tavalla hakijan ominaisuudet vastasivat asetettuja kriteereitä. Toisaalta valintakriteerien avaaminen osittain työnhakijoille ennen hakemuksen jättämistä on myös vaihtoehtona, esimerkiksi tärkeimmistä painotusalueista, mutta tässä suunnitteluprosessissa ne huomioidaan hakemuksen jättämisen jälkeisissä rekrytointivaiheissa, koska pyritään estämään väärinkäyttöjen mahdollisuutta.

Tulosten vertailutietoon liittyvien tulosten perusteella haasteltavat toivoivat rekrytointiprosessin loppuvaiheisiin. Tämä on myös loogista, sillä ne ovat vaiheita, joissa hakijoiden ansioluetteloita on jo analysoitu sekä seulottu ja tehty mahdollisesti työtehtävään soveltavia testejä heille. Esimerkiksi koonti rekrytointiin liittyvistä vertailutiedoista voisi toimia palauteviestin yhteydessä. Toisaalta työnhakijat toivovat myös reaaliaikaista ja automaattista tiedottamista omiin rekrytointituloksiin liittyen, joten esimerkiksi tietynlainen liveseuranta olisi myös tärkeää. Tämä voisi olla tilanteessa, jossa tekoälypohjainen rekrytointijärjestelmä automaattisesti analysoi ansioluetteloita sekä saatekirjeitä ja näyttää vertailutietoja henkilön hakemuksen tilanteesta sen hetkisestä tilanteesta.

Kehittävän palautteen suhteen rekrytointiprosessin loppuvaiheessa voi tarjota työnhakijalle palautteen, joka auttaa ymmärtämään, miten heidän osaamisensa ja kokemuksensa ovat suhteutuneet työn vaatimuksiin ja muihin hakijoihin. Tämä

palautetieto esitetään yleisluontoisesti siten, että se antaa hakijalle käsityksen hänen vahvuuksistaan ja mahdollisista kehityskohteista ilman, että tietosuoja vaarantuu tai muiden hakijoiden tulokset paljastuvat. Palautteessa voitaisiin esimerkiksi tuoda, millä osa-alueilla työnhakijan osaaminen oli linjassa työn keskeisten kriteerien kanssa, ja missä kohdissa hän voisi vahvistaa osaamistaan tulevia hakuja varten.

Reaaliaikaisesta ja automaattisesta tiedottamisesta käyttötapaus voisi olla teorian perusteella tilanne, jossa analysoidaan työnhakijan ansioluettelon koneluettavuus. Tällainen käyttötapaus tukee työnhakijoiden toivetta saada nopeaa palautetta hakemuksen eri osa-alueista ja osoittaa, miten järjestelmä voi tarjota hakijalle automaattisia ilmoituksia hakemuksen tilanteesta. Ansioluettelon koneluettavuuden analysointi antaa työnhakijalle välitöntä palautetta siitä, kuinka hyvin heidän ansioluettelonsa täyttää järjestelmän määrittämät kriteerit.

Mahdollisuus kysyä lisätietoja käyttötapauksena voisi olla tilanne, jossa työnhakija voi esittää kysymyksiä tekoälypohjaiselle rekrytointijärjestelmälle, esimerkiksi chatbotin kautta. Tämä käyttötapaus vastaa työnhakijoiden tarpeeseen saada syvällisempää tietoa ja perusteluja hakemusten käsittelyyn liittyvistä päätöksistä, erityisesti hylkäystilanteissa. Chatbot voi tarjota hakijalle tarkempia selityksiä siitä, mihin päätökset perustuvat, sekä mahdollisuuden esittää lisäkysymyksiä, jos hakija kaipaa vielä syvempää ymmärrystä. Tämä vuorovaikutteinen ominaisuus lisää rekrytointiprosessin läpinäkyvyyttä ja tukee hakijan tiedonsaantia, mikä puolestaan vahvistaa luottamusta tekoälypohjaisen järjestelmän toimintaan.

Automaatio ja ihmisten roolien selkeyttäminen käyttötapauksena tarkoittaa vastuualueiden jakamista ja viestintää eri rekrytointiprosessin vaiheissa. Tämä selkeyttää työnhakijalle, missä vaiheissa tekoäly osallistuu prosessiin ja milloin inhimillinen arviointi astuu kuvaan. Tällainen läpinäkyvyys tukee hakijan ymmärrystä rekrytointiprosessin etenemisestä ja lisää luottamusta päätöksentekoon. Esimerkiksi hakemuksen jättämisen yhteydessä tekoäly voi tarkistaa koneluettavuuden ja täydentämättömät tiedot.

Seulontavaiheessa tekoäly analysoi hakemuksia määriteltyjen kriteerien pohjalta, kun taas arviointivaiheessa rekrytoija voi tarkastella seulonnan läpäisseitä hakemuksia. Lopullisessa vaiheessa rekrytoija suorittaa haastattelut ja valinnan, jolloin ihmisen rooli on keskeinen. Palautteen antamisessa tekoäly voi muodostaa palautetta, ja rekrytoija voi antaa loppuvaiheessa henkilökohtaista palautetta.

Seuraavaksi suunnitteluprosessissa muodostetaan selitysmallit ja selittävä informaatio käyttötapauksille. Tässä tutkielmassa keskitytään konseptuaaliseen suunnitteluun, eli mallinnetaan, miten tekoälypohjainen rekrytointijärjestelmä voi tarjota läpinäkyvää ja ymmärrettävää tietoa käyttäjilleen. Selitysmallit voivat sisältää esimerkiksi päätöksenteon perusteet, tärkeimmät valintakriteerit sekä automaation ja ihmisen roolit rekrytointiprosessin eri vaiheissa.

Tämän jälkeen muodostui ohjeistuksen rakenne, joka keskittyy selitettävyyden ja vuorovaikutuksellisuuden konseptuaaliseen kehittämiseen tekoälypohjaisessa rekrytointijärjestelmässä. Ohjeistuksen tarkoituksena on selkeyttää, miten tekoäly voi tarjota käyttäjille ymmärrettävää ja läpinäkyvää tietoa päätöksenteon eri vaiheissa sekä tukea hakijoiden tiedontarpeita prosessin aikana. Ohjeistuksen rakenne keskittyy selitettävyyden ja vuorovaikutuksellisuuden konseptuaaliseen kehittämiseen tekoälypohjaisessa rekrytointijärjestelmässä. Sen tarkoituksena on selkeyttää, miten tekoäly voi tarjota käyttäjille ymmärrettävää ja läpinäkyvää tietoa päätöksenteon eri vaiheissa sekä tukea hakijoiden tiedontarpeita prosessin aikana.

Ohjeistuksessa esitetään ensin suunnittelumalli, jossa kehitetään selitysmalleja ja selittävää informaatiota käyttötapauksille. Tässä osiossa määritellään selitettävyyden ja vuorovaikutuksellisuuden konseptit rekrytointiprosessin eri vaiheisiin, kuten hakemuksen jättämiseen, seulontaan, arviointiin ja palautteen antamiseen. Tämän jälkeen tutkimuksen tuloksiin perustuvat visuaaliset esimerkit havainnollistavat selitysmallien käytännön toteutusta. Esimerkkeinä ovat reaaliaikainen tiedottaminen, vastuiden selkeyttäminen automaation ja ihmisen roolien välillä sekä mahdollisuus kysyä

lisätietoja päätöksistä chatbotin kautta. Näiden visuaalisten esimerkkien avulla työnhakijat voivat ymmärtää selityksien sisällön ja prosessin logiikan selkeämmin.

Ohjeistuksen rakenteen tarkoitus on tukea käyttäjiä tarjoamalla kattavan käsityksen siitä, miten tekoälypohjainen rekrytointijärjestelmä voisi toimia ja miten vuorovaikutus eri vaiheissa toteutuisi. Lisäksi tarkoitus on auttaa rekrytointin kehittäjiä suunnittelemaan ja ideoimaan ratkaisuja, jotka vastaavat työnhakijoiden selitettävyyden ja vuorovaikutustarpeisiin.

6.3 Ohjeistuksen toteutus

Ohjeistuksen toteuttamistavaksi valikoitui infografiikka, sillä mahdollistaa tiedon esittämisen tiiviissä muodossa. Infografiikan toteutuksessa käytettiin Canva-sovellusta, jonka avulla luotiin visuaalisia esimerkkejä rekrytointiprosessin selitettävyydestä ja vuorovaikutuksellisuudesta. Canva mahdollisti selkeiden ja intuitiivisten elementtien, kuten pisteytysten, vertailutulosten ja reaaliaikaisen palautteen visualisoinnin, joiden avulla työnhakijat voivat saada selkeämmän käsityksen tekoälypohjaisen rekrytointiprosessin logiikasta ja omasta sijoittumisestaan rekrytointikriteereihin. Infografiikan suunnittelussa otettiin huomioon myös vuorovaikutukselliset tarpeet, kuten mahdollisuus kysyä lisätietoja sekä automaation ja ihmisten roolien selkeyttäminen prosessin eri vaiheissa. Lopullinen infografiikka, joka havainnollistaa näitä käytännön sovelluksia, on liitteenä.

Infografiikka rakentuu kolmen pääteeman ympärille, joita ovat selitettävyyden ja vuorovaikutuksellisuuden konseptuaalinen kehittäminen, selitettävyyden näkökulmat ja vuorovaikutuksellisuuden kehityssuunnat. Jokainen näistä teemoista tukee työnhakijoiden kokemusta tekoälypohjaisessa rekrytointiprosessissa. Infografiikan toteutus tuo esiin, miten selitettävyys ja vuorovaikutus on muodostettu vastaamaan työnhakijoiden tarpeita, ja havainnollistaa tekoälyn tarjoamien tietojen merkityksen eri vaiheissa.

Ensimmäinen pääteema, selitettävyyden ja vuorovaikutuksellisuuden konseptuaalinen kehittäminen, keskittyy tekoälypohjaisen rekrytointijärjestelmän toiminnan periaatteisiin ja käyttötapauksiin. Tässä osiossa määritellään selitettävyyden ja vuorovaikutuksellisuuden konseptit, kuten selitysvaatimukset, kontekstikuvaukset ja vuorovaikutustarpeet, jotka ovat olennaisia rekrytointiprosessin eri vaiheissa. Konseptuaalisella kehittämisellä pyritään luomaan perusta käyttäjien tiedontarpeiden täyttämiseksi ja järjestelmän käytön ymmärrettävyydelle.

Toinen pääteema, selitettävyyden näkökulmat, keskittyy konkreettisten valintakriteerien avaamiseen, mikä tarjoaa työnhakijoille näkyvyyttä siihen, miten heidän tietonsa vastaavat valintaprosessin kriteereitä ja mitkä tekijät järjestelmä huomioi arvioinnissaan. Tämä on esitetty visuaalisesti selkeillä kaavioilla ja taulukoilla, jotka kuvaavat pisteytysperiaatteita ja selventävät, mitkä tiedot vaikuttavat lopullisiin arviointeihin. Tulosten vertailutiedot muodostavat selitettävyysoSION toisen osan, jossa työnhakijalle tarjotaan tietoa omasta sijoittumisestaan suhteessa muihin hakijoihin. Näin hakija voi hahmottaa omaa osaamistasoaan ja tunnistaa mahdollisia kehityskohteita. Osion lopuksi käsitellään palautteen tarjoamista hakijan kehityksen tueksi ja prosessiperusteista läpinäkyvyyttä, mikä auttaa työnhakijaa ymmärtämään tekoälypohjaisen päätöksenteon kulun ja valintaperusteet rekrytointiprosessin edetessä.

Kolmas pääteema, vuorovaikutuksellisuuden kehityssuunnat, painottaa reaaliaikaista ja automaattista tiedottamista, jonka avulla työnhakijalle tarjotaan ajantasaista tietoa hänen hakemuksensa etenemisestä ja päätöksenteon eri vaiheista. Tämä mahdollistaa prosessin seuraamisen ja valmistautumisen seuraaviin vaiheisiin. Vuorovaikutuksellisuuden kehityssuunnat sisältävät myös hakijalle tarjottavan mahdollisuuden kysyä lisätietoja, erityisesti tilanteissa, joissa hakemuksen eteneminen päättyy. Näin hakija voi saada selvennyksiä päätöksen perusteista ja hyödyntää tietoa tulevissa hakukierroksissa. Lopuksi tämä teema käsittelee automaation ja ihmisten roolien selkeyttämistä rekrytointiprosessissa, jossa infografiikka havainnollistaa, missä kohdin tekoäly tekee itsenäisiä arvioita ja milloin ihmisen osallistuminen on ratkaisevaa.

Tämä jaottelu tuo prosessiin läpinäkyvyyttä, jolloin työnhakija voi ymmärtää, missä vaiheissa hänen tietojaan käsittelee tekoäly ja milloin rekrytoijan arviointi vaikuttaa päätöksiin. Infografiikan rakenne tukee näin ollen työnhakijoiden tarpeita ymmärtää prosessin logiikkaa, selitettävyyksiperiaatteita ja vuorovaikutuksellisia mahdollisuuksia, joita tekoälypohjainen rekrytointi voi tarjota.

6.4 Arviointi

Haastateltavien näkökulmasta infografiikka edustaa hyvin heidän toiveitaan selitettävyydestä tekoälypohjaisessa rekrytoinnissa, sillä se tarjoaa selkeän yleiskuvan prosessin kulusta ja eri valintakriteerien merkityksestä. Visuaaliset esitystavat, kuten pisteytystaulukot ja graafiset kuvaajat, vahvistavat työnhakijoiden ymmärrystä siitä, miten heidän tietojaan käsitellään ja miten eri osa-alueet, kuten koulutus ja työkokemus, vaikuttavat heidän sijoittumiseensa. Tämä visuaalinen esitystapa vastaa hakijoiden toiveita selkeyden lisäämisestä prosessiin ja auttaa heitä hahmottamaan tekoälyn tekemien arviointien perusteet. Useat haastateltavat kokivat, että konkreettisten valintakriteerien avaaminen tukee heidän tarvettaan ymmärtää, mitä hakuprosessissa arvostetaan ja mitä tietoja tekoäly analysoi arvioinnin pohjaksi.

Infografiikka onnistuu myös vastaamaan työnhakijoiden odotuksiin siitä, miten tekoäly voisi tarjota heille hyödyllistä ja kehittävää tietoa omasta suoriutumisestaan. Esittämällä kriteerikohtaiset pisteytykset ja vertailutiedot infografiikka auttaa työnhakijoita arvioimaan omia vahvuuksiaan ja tunnistamaan kehityskohteitaan. Haastateltavat pitivät infografiikkaa käytännöllisenä työkaluna, sillä se selkeyttää heidän käsitystään omasta sijoittumisestaan rekrytointiprosessissa ja mahdollistaa oman osaamisen suhteuttamisen muihin hakijoihin. Useat haastateltavat näkivätkin infografiikan hyödyllisenä tapana esittää analyysitiedot selkeästi ja ymmärrettävästi, mikä tukee heidän toivettaan siitä, että tekoälypohjainen rekrytointi voisi tarjota paitsi seulontatyökalun myös välineen, joka auttaa oppimaan omista mahdollisuuksista ja valmistautumaan paremmin tuleviin hakemuksiin.

Vuorovaikutuksellisuuden kannalta infografiikka edustaa työnhakijoiden toiveita selkeän viestinnän ja vuorovaikutteisten mahdollisuuksien lisäämiseksi. Työnhakijat arvostivat erityisesti infografiikan tapaa esittää reaaliaikaisen tiedottamisen mahdollisuudet ja tarjota konkreettisia esimerkkejä siitä, missä prosessin vaiheissa he voivat odottaa palautetta tai lisätietoja. Tämä vastaa heidän toivettaan avoimemmasta ja vuorovaikutteisemmin etenevästä rekrytointiprosessista, jossa automaation tuottamat arviot ja tiedot ovat työnhakijan saatavilla. Näin he kokevat voivansa seurata oman hakemuksensa edistymistä ja oppia paremmin tekoälyn arviointiperusteista. Mahdollisuus kysyä lisätietoja prosessin päättyessä koettiin myös arvokkaaksi, sillä se auttaa hakijoita ymmärtämään, missä vaiheessa heidän suoriutumisen jää mahdollisesti heikommaksi ja mitä he voisivat parantaa tulevissa haussa.

Työnhakijat näkivät infografiikan hyödyllisenä myös siksi, että se erittelee automaation ja ihmisten roolit prosessin eri vaiheissa. Tämä koettiin tärkeäksi vuorovaikutuksellisuuden ja selitettävyyden kannalta, sillä se antaa työnhakijoille mahdollisuuden ymmärtää, missä kohtaa ihmisen tekemä arvio vaikuttaa päätöksiin ja milloin tekoäly toimii itsenäisesti. Infografiikka vastaa heidän odotuksiinsa selkeyttämällä näitä rooleja ja osoittamalla, millä tavoin tekoäly ja ihmiset toimivat yhdessä tukemassa päätöksentekoa. Tämä selkeytys tukee heidän toivettaan läpinäkyvämmästä ja oikeudenmukaisemmasta prosessista, jossa automaation vaikutus on näkyvää ja ymmärrettävää.

Haastateltavat toivat esiin myös kehitysideoita, joilla infografiikka voisi palvella työnhakijoita vielä paremmin. Ensinnäkin osa työnhakijoista toivoi tarkempaa tietoa valintakriteerien painotuksista, kuten siitä, kuinka paljon eri osa-alueet – esimerkiksi koulutus tai työkokemus – vaikuttavat lopulliseen pisteytykseen. He kokivat, että tällainen yksityiskohtaisempi tieto lisäisi ymmärrystä siitä, mitä taitoja ja kokemuksia painotetaan erityisesti juuri kyseisessä rekrytoinnissa ja auttaisi hakijoita valmistautumaan ja kehittämään oikeita osaamisalueita.

Toinen kehitysehdotus liittyi palautteen yksilöllistämiseen ja tarkempien vertailutietojen tarjoamiseen. Haastateltavat mainitsivat, että olisi hyödyllistä saada esimerkkien lisäksi tarkempaa tietoa siitä, miten oma suoriutuminen suhteutuu keskimääräisiin pisteisiin tai parhaimpiin hakijoihin nähden. Tämä tukisi työnhakijoiden toivetta saada konkreettisempaa palautetta omista kehitysalueistaan ja nähdä tarkemmin, missä suhteessa heidän taitonsa ovat muihin hakijoihin. Yksityiskohtaisempi palaute tukisi heidän valmiuksiaan parantaa tulevia hakemuksiaan ja keskittyä kehittämään juuri niitä osaamisalueita, jotka ovat kriittisiä rekrytointimenestyksen kannalta.

Vuorovaikutuksellisuuden osalta työnhakijat toivoivat lisää mahdollisuuksia saada reaaliaikaista palautetta rekrytointiprosessin edetessä. Erityisesti hakuprosessin alkuvaiheessa automaattisesti tuotettu palaute voisi auttaa hakijoita korjaamaan mahdollisia puutteita esimerkiksi CV sisällössä tai täyttämään puuttuvia tietoja jo hakuvaiheessa. Tällainen ennakoiva palaute voisi myös auttaa heitä ymmärtämään tekoälyn käsittelyvaiheita paremmin ja varmistamaan, että heidän hakemuksensa vastaa järjestelmän vaatimuksiin.

Lopuksi haastateltavat ehdottivat, että infografiikka voisi selittää vielä tarkemmin tekoälyn ja ihmisten roolien yhteispeliä rekrytointiprosessin eri vaiheissa. Vaikka infografiikka selkeytti automaation ja ihmisten osuutta päätöksenteossa, osa työnhakijoista koki, että yksityiskohtaisempi selitys siitä, missä ja miten rekrytoija itse osallistuu prosessiin, voisi lisätä luottamusta järjestelmään. He uskovat, että tällainen läpinäkyvyys auttaisi työnhakijoita ymmärtämään paremmin, missä vaiheessa ihmisen tekemä arviointi voi täydentää tekoälyn tekemää analyysia, mikä lisäisi heidän kokemustaan prosessin oikeudenmukaisuudesta ja inhimillisyydestä.

Näiden näkökulmien ja kehitysehdotusten pohjalta infografiikassa olevat esimerkit voisivat parantaa työnhakijoiden kokemusta ja tarjota heille entistä syvällisempää ymmärrystä tekoälypohjaisen rekrytointin toimintalogiikasta ja vuorovaikutteisuudesta. Infografiikan kehitysviesti vastaa työnhakijoiden odotuksia rekrytointijärjestelmien

kehittäjiltä, sillä työnhakijat arvostavat läpinäkyvyyttä ja konkreettisia tietoja siitä, miten heidän suoriutumisensa arvioidaan sekä millä perusteilla päätöksiä tehdään. Lisäksi he odottavat vuorovaikutteisuutta ja mahdollisuutta ymmärtää ja vaikuttaa omaan sijoittumiseensa prosessin eri vaiheissa. Infografiikka tarjoaa kehittäjille käsitteellisiä ohjeita siihen, miten näitä odotuksia voidaan huomioida ja täyttää, jotta rekrytointiprosessi olisi mahdollisimman avoin, oikeudenmukainen ja työnhakijoiden kannalta mielekäs.

7 Johtopäätökset ja pohdinta

Tutkimus selitettävyydestä ja vuorovaikutuksellisuudesta tekoälypohjaisissa rekrytointijärjestelmissä osoitti, että työnhakijoilla on vahvoja odotuksia läpinäkyvyyden ja ymmärrettävyyden suhteen rekrytointiprosessin aikana. Aiemmat tutkimukset, kuten Hunkenschroer & Luetge (2022), Langer & König (2023) sekä Leslie & Briggs (2021), ovat jo osoittaneet, että työnhakijat haluavat selkeitä ja ymmärrettäviä selityksiä tekoälypohjaisten järjestelmien päätöksille sekä mahdollisuuden vuorovaikutukseen järjestelmän kanssa, esimerkiksi kysymällä lisätietoja tai pyytämällä tarkennuksia. Tämä tutkimus vahvistaa työnhakijoiden odotukset selkeistä prosessi- ja tulosperusteisista selityksistä ja tukee aikaisempien tutkimusten havaintoja vuorovaikutuksellisuuden merkityksestä.

Löydökset osoittavat, että selitettävyyksivaatimukset rekrytointijärjestelmissä voidaan täyttää tarjoamalla visuaalisesti ja sanallisesti helposti ymmärrettäviä selityksiä tekoälyn päätöksille. Koivusen ja muiden (2022) tutkimus vahvisti, että reaaliaikainen tiedon tarjoaminen hakijoille heidän tilanteestaan prosessin eri vaiheissa voi lisätä luottamusta ja kokemusta avoimuudesta. Myös Haresamudram ja muiden (2023) mukaan läpinäkyvyys ja vuorovaikutusmahdollisuuksien lisääminen ovat keskeisiä hakijakokemuksen parantamiseksi, ja nämä havainnot tukevat myös tässä tutkimuksessa muodostettua ohjeistusta, joka auttaa rekrytoinnin kehittäjiä vastaamaan työnhakijoiden toiveisiin selitettävyydestä ja vuorovaikutuksesta. Lisäksi tekoälyn ja ihmisen roolien selkeyttäminen rekrytointiprosessissa on keskeistä, jotta työnhakijat ymmärtävät, kuinka sekä tekoäly että ihmiset vaikuttavat heidän arviointiinsa (Langer & König, 2023; Leslie & Briggs, 2021).

Lisäksi järjestelmien jatkuva arviointi ja mukauttaminen on olennaista tekoälyn kehittämisessä. Hunkenschroer & Luetge (2022) painottavat, että järjestelmien tulee olla joustavia ja niillä tulee olla kyky oppia käyttäjiltä saadusta palautteesta. Tekoälyjärjestelmien tulee siis mukauttaa toimintaansa työnhakijoilta saadun palautteen perusteella, jotta ne voivat paremmin vastata muuttuvia odotuksia ja tarpeita, mikä

parantaisi järjestelmän oikeudenmukaisuutta ja hakijakokemusta pitkällä aikavälillä. Tämä on vuorovaikutusparadigma, joka mahdollisesti lisää hakijoiden luottamusta järjestelmään ja kannustaa heitä osallistumaan aktiivisesti rekrytointiprosessiin.

Hakijoiden vaikutusmahdollisuuksien lisääminen päätöksentekoprosessissa on myös tärkeä kehityssuunta. Työnhakijat haluavat mahdollisuuden vaikuttaa ja saada tarkempia tietoja siitä, miten päätöksiä tehdään ja miten heidän tietojaan käytetään (Koivunen ym., 2022). Esimerkiksi keskusteleva käyttöliittymä tai chattibotti voi tarjota hakijoille tilaisuuden esittää kysymyksiä ja saada selkeitä vastauksia päätöksentekoon liittyvistä yksityiskohdista. Tällainen toiminnallisuus voisi lisätä hakijoiden tyytyväisyyttä ja luottamusta järjestelmään, sillä he tuntisivat voivansa aktiivisesti vaikuttaa prosessiin ja ymmärtää sen perusteita.

Vaikka tutkimus tarjoaa tärkeitä näkökulmia selitettävyyden ja vuorovaikutuksellisuuden kehittämiseksi tekoälypohjaisissa rekrytointijärjestelmissä, siihen liittyy myös ongelmia ja jatkotutkimuksen tarpeita. Felzmann ja muut (2020) tuovat esiin, että selitettävyyslähestymistapojen monimutkaisuus asettaa ongelmia erityisesti korkean teknologian järjestelmissä, joissa päätöksentekoprosessit voivat olla vaikeasti ymmärrettäviä niin työnhakijoille kuin kehittäjille. Tämä asettaa paineita kehittää käyttäjäystävällisiä ja intuitiivisia käyttöliittymiä, jotka mahdollistavat monimutkaisen teknologian selittämisen yksinkertaisella tavalla.

Toinen keskeinen pohdinnan kohde on tekoälyjärjestelmien mahdollinen puolueellisuus, joka voi vaikuttaa työnhakijoiden tasa-arvoiseen kohteluun. Esimerkiksi Van Iddekingen ja muiden (2023) sekä Hunkenschroerin & Luetgen (2022) tutkimukset ovat osoittaneet, että eettisiin näkökohtiin on tärkeää kiinnittää huomiota, sillä tiedon automaatio ja analytiikka voivat vaikuttaa eri työnhakijaryhmiin eri tavoin. Aiemmat tapaukset, kuten Amazonin järjestelmän syrjivät piirteet Haresamudram ja muut (2023), korostavat eettisten periaatteiden merkitystä, ja tämän tutkimuksen havaintoihin perustuen tulevaisuuden tutkimus voisi keskittyä selkeämmin myös tiettyjen työnhakijaryhmien,

kuten eri ikäryhmien odotuksiin ja kokemuksiin tekoälypohjaisista järjestelmistä (Albassam, 2023).

Jatkotutkimusta tarvitaan myös arvioimaan, miten tässä tutkimuksessa kehitetty konseptuaalinen ohjeistus vaikuttaa käytännössä rekrytointiprosessien ja -järjestelmien kehittämiseen. Esimerkiksi Langer & König (2023) sekä Haresamudram ja muut (2023) ehdottavat pilottiohjelmien hyödyntämistä, jotta tekoälyjärjestelmien vaikutuksia hakijakokemukseen ja sitoutumiseen voidaan seurata. Tällainen empiirinen arviointi tukisi tämän tutkimuksen ohjeistuksen jatkokehitystä ja sen muokkaamista entistä paremmin työnhakijoiden toiveita vastaavaksi.

Lähteet

- AI Pact | Shaping Europe's digital future*. (ei pvm.). Noudettu 21. syyskuuta 2024, osoitteesta <https://digital-strategy.ec.europa.eu/en/policies/ai-pact>
- Alasuutari, P. (2011). *Laadullinen tutkimus 2.0*. Vastapaino.
- Albaroudi, E., Mansouri, T., & Alameer, A. (2024). A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring. *AI*, 5(1), Article 1. <https://doi.org/10.3390/ai5010019>
- Albassam, W. A. (2023). The Power of Artificial Intelligence in Recruitment: An Analytical Review of Current AI-Based Recruitment Strategies. *International Journal of Professional Business Review*, 8(6), e02089. <https://doi.org/10.26668/businessreview/2023.v8i6.2089>
- Alexander Jung. (2022). *Machine Learning*. Springer.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Benyon, D. (2019). *Designing user experience: A guide to HCI, UX and interaction design* (Fourth edition). Pearson Education Limited.
- Bernardo, E., & Rosemary, S. (2023). Affective Design Analysis of Explainable Artificial Intelligence (XAI): A User-Centric Perspective. *Informatics (Basel)*, 10(1), 32-. <https://doi.org/10.3390/informatics10010032>

- Bingley, W. J., Curtis, C., Lockey, S., Bialkowski, A., Gillespie, N., Haslam, S. A., Ko, R. K. L., Steffens, N., Wiles, J., & Worthy, P. (2023). Where is the human in human-centered AI? Insights from developer priorities and user experiences. *Computers in Human Behavior, 141*, 107617. <https://doi.org/10.1016/j.chb.2022.107617>
- Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). "It's Reducing a Human Being to a Percentage": Perceptions of Justice in Algorithmic Decisions. <https://doi.org/10.1145/3173574.3173951>
- Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons, 63*(2), 215–226. <https://doi.org/10.1016/j.bushor.2019.12.001>
- Bovens, M. A. P., Goodin, R. E., & Schillemans, T. (2014). *The Oxford handbook of public accountability*. Oxford University Press.
- Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Boselie, P., Lee Cooke, F., Decker, S., DeNisi, A., Dey, P. K., Guest, D., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., ... Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal, 33*(3), 606–659. <https://doi.org/10.1111/1748-8583.12524>
- Cálem, J., Moreira, C., & Jorge, J. (2024). Intelligent systems in healthcare: A systematic survey of explainable user interfaces. *Computers in Biology and Medicine, 180*, 108908. <https://doi.org/10.1016/j.combiomed.2024.108908>
- Campeato, O. (2020). *Artificial Intelligence, Machine Learning, and Deep Learning*. Mercury Learning & Information. https://sfx.finna.fi/nelli07?url_ver=Z39.88-

2004&ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rfr_id=info:sid/sfxit.com:opac_856&url_ctx_fmt=info:ofi/fmt:kev:mtx:ctx&sfx.ignore_date_threshold=1&rft.object_id=4100000010135200&svc_val_fmt=info:ofi/fmt:kev:mtx:sch_svc&

Chromik, M., & Butz, A. (2021). *Human-XAI Interaction: A Review and Design Principles for Explanation User Interfaces*. https://doi.org/10.1007/978-3-030-85616-8_36

Clement, T., Kemmerzell, N., Abdelaal, M., & Amberg, M. (2023). XAIR: A Systematic Metareview of Explainable AI (XAI) Aligned to the Software Development Process. *Machine Learning and Knowledge Extraction*, 5(1), 78–108. <https://doi.org/10.3390/make5010006>

Conte, F., & Siano, A. (2023). Data-driven human resource and data-driven talent management in internal and recruitment communication strategies: An empirical survey on Italian firms and insights for European context. *Corporate Communications*, 28(4), 618–637. <https://doi.org/10.1108/CCIJ-02-2022-0012>

Dahlbom, P., Siikanen, N., Sajasalo, P., & Järvenpää, M. (2020). Big data and HR analytics in the digital era. *Baltic Journal of Management*, 15(1), 120–138. <https://doi.org/10.1108/BJM-11-2018-0393>

Dreamson, N., & Khine, P. H. H. (2022). Abductive Reasoning: A Design Thinking Experiment. *International Journal of Art & Design Education*, 41(3), 403–413. <https://doi.org/10.1111/jade.12424>

Duong, N. T., & Thi, T. D. P. (2022). HOW DOES AI RECRUITMENT INFLUENCE SATISFACTION AMONG STUDENT JOB-SEEKERS? THE ROLE OF SELF-EFFICACY AS

A MODERATOR AND MEDIATOR. *Obrazovanie i Nauka*, 24(8), 64–94.

<https://doi.org/10.17853/1994-5639-2022-8-64-94>

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T.,

Morgan, G., & Ranjan, R. (2023). Explainable AI (XAI): Core Ideas, Techniques, and

Solutions. *ACM Computing Surveys*, 55(9), 1–33.

<https://doi.org/10.1145/3561048>

Ehsan, U., & Riedl, M. (2020). Human-centered Explainable AI: Towards a Reflective

Sociotechnical Approach. *arXiv.Org*. <https://doi.org/10.48550/arxiv.2002.01092>

Ehsan, U., Tambwekar, P., Chan, L., Harrison, B., & Riedl, M. (2019). *Automated rationale*

generation: A technique for explainable AI and its effects on human perceptions.

<https://doi.org/10.1145/3301275.3302316>

Faliagka, E., Tsakalidis, A., & Tzimas, G. (2012). An integrated e-recruitment system for

automated personality mining and applicant ranking. *Internet Research*, 22(5),

551–568. <https://doi.org/10.1108/10662241211271545>

Fan, H., Han, B., & Gao, W. (2022). (Im)Balanced customer-oriented behaviors and AI

chatbots' Efficiency–Flexibility performance: The moderating role of customers'

rational choices. *Journal of Retailing and Consumer Services*, 66, 102937–.

<https://doi.org/10.1016/j.jretconser.2022.102937>

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrioux, A. (2020). Towards

Transparency by Design for Artificial Intelligence. *Science and Engineering Ethics*,

26(6), 3333–3361. <https://doi.org/10.1007/s11948-020-00276-4>

- Gao, Z. (2021). Precise position intelligent matching system of online recruitment platform based on data mining technology. *Journal of Physics: Conference Series*, 2066(1). <https://doi.org/10.1088/1742-6596/2066/1/012001>
- Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, 21(2), 135–146. <https://doi.org/10.1057/ejis.2011.54>
- Goldkuhl, G. (2019). The generation of qualitative data in information systems research: The diversity of empirical research methods. *Communications of the Association for Information Systems*, 44, 572–599.
- Goodman, B., & Flaxman, S. (2017). European Union Regulations on Algorithmic Decision Making and a “Right to Explanation”. *The AI Magazine*, 38(3), 50–57. <https://doi.org/10.1609/aimag.v38i3.2741>
- Guidotti, R., Monreala, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedrechi, D. (2019). A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, 51(5), 1–42. <https://doi.org/10.1145/3236009>
- Gunning, D., & Aha, D. W. (2019). DARPA’s Explainable Artificial Intelligence Program. *The AI Magazine*, 40(2), 44–58. <https://doi.org/10.1609/aimag.v40i2.2850>
- Haresamudram, K., Larsson, S., & Heintz, F. (2023). Three Levels of AI Transparency. *Computer (Long Beach, Calif.)*, 56(2), 93–100. <https://doi.org/10.1109/MC.2022.3213181>
- Hoffman, R. R., & Klein, G. (2017). Explaining Explanation, Part 1: Theoretical Foundations. *IEEE Intelligent Systems*, 32(3), 68–73. <https://doi.org/10.1109/MIS.2017.54>

- Hollanek, T. (2023). AI transparency: A matter of reconciling design with critique. *AI & Society, 38*(5), 2071–2079. <https://doi.org/10.1007/s00146-020-01110-y>
- Holm, A. B. (2012). E-recruitment: Towards an Ubiquitous Recruitment Process and Candidate Relationship Management. *ZEITSCHRIFT FÜR PERSONALFORSCHUNG, 26*(3), 241–259. https://doi.org/10.1688/1862-0000_ZfP_2012_03_Holm
- Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda. *Journal of Business Ethics, 178*(4), 977–1007. <https://doi.org/10.1007/s10551-022-05049-6>
- Janiesch, C. (2021). Machine learning and deep learning. *Electronic Markets, 31*(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Johnson, R. D., Stone, D. L., & Lukaszewski, K. M. (2021). The benefits of eHRM and AI for talent acquisition. *Journal of Tourism Futures, 7*(1), 40–52. <https://doi.org/10.1108/JTF-02-2020-0013>
- Kasanen, E., Lukka, K., & Siitonen, A. (1993). The constructive approach in management accounting research. *Journal of Management Accounting Research, 5*, 243.
- Koivunen, S., Ala-Luopa, S., Olsson, T., & Haapakorpi, A. (2022). The March of Chatbots into Recruitment: Recruiters' Experiences, Expectations, and Design Opportunities. *Computer Supported Cooperative Work, 31*(3), 487–516. <https://doi.org/10.1007/s10606-022-09429-4>
- Köchling, A., Wehner, M. C., & Warkocz, J. (2023). Can I show my skills? Affective responses to artificial intelligence in the recruitment process. *Review of Managerial Science, 17*(6), 2109. <https://doi.org/10.1007/s11846-021-00514-4>

- Langer, M., & König, C. J. (2023). Introducing a multi-stakeholder perspective on opacity, transparency and strategies to reduce opacity in algorithm-based human resource management. *Human Resource Management Review*, 33(1), 100881. <https://doi.org/10.1016/j.hrmr.2021.100881>
- Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, 25(4), 371–382. <https://doi.org/10.1111/ijsa.12191>
- Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sasing, A., & Baum, K. (2021). What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence*, 296, 103473. <https://doi.org/10.1016/j.artint.2021.103473>
- Laumer, S., Maier, C., & Eckhardt, A. (2015). The impact of business process management and applicant tracking systems on recruiting process performance: An empirical study. *Zeitschrift Für Betriebswirtschaft*, 85(4), 421–453. <https://doi.org/10.1007/s11573-014-0758-9>
- Lee, I. (2007). An architecture for a next-generation holistic e-recruiting system. *Communications of the ACM*, 50(7), 81–85. <https://doi.org/10.1145/1272516.1272518>
- Leslie, D., & Briggs, M. (2021). Explaining decisions made with AI: A workbook (Use case 1: AI-assisted recruitment tool). *arXiv.Org*. <https://doi.org/10.48550/arxiv.2104.03906>

- Lookadoo, K., & Moore, S. (2024). Is Your Résumé/Textbook Up-To-Date? An Audit of AI ATS Résumé Instruction. *Business and Professional Communication Quarterly*.
<https://doi.org/10.1177/23294906231223101>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Nikolaou, I. (2021). What is the Role of Technology in Recruitment and Selection? *The Spanish Journal of Psychology*, 24, e2. <https://doi.org/10.1017/SJP.2021.6>
- Norman, P. (2022). *HR*. Knowledge Resources.
- Osaba, E. (2021). *Artificial Intelligence*. IntechOpen.
- Palshikar, G. K., Pawar, S., Banerjee, A. S., Srivastava, R., Ramrakhiyani, N., Patil, S., Thosar, D., Bhat, J., Jain, A., Hingmire, S., Chaurasia, S., Mandloi, P., & Chalavadi, D. (2023). RINX: A system for information and knowledge extraction from resumes. *Data & Knowledge Engineering*, 147, 102202.
<https://doi.org/10.1016/j.datak.2023.102202>
- Pillai, R., Yadav, S., Sivathanu, B., Kaushik, N., & Goel, P. (2022). Use of 4.0 (I4.0) technology in HRM: a pathway toward SHRM 4.0 and HR performance. *Foresight (Cambridge)*, 24(6), 708–727. <https://doi.org/10.1108/FS-06-2021-0128>
- Puusa, A., Juuti, P., & Aaltio, I. (Toim.). (2020). *Laadullisen tutkimuksen näkökulmat ja menetelmät*. Gaudeamus.
- Rego, A., Cunha, M. P. e, & Giustiniano, L. (2022). Are Relationally Transparent Leaders More Receptive to the Relational Transparency of Others? An Authentic Dialog Perspective. *Journal of Business Ethics*, 180(2), 695–709.
<https://doi.org/10.1007/s10551-021-04792-6>

- Ridley, M. (2022). Explainable Artificial Intelligence (XAI): Adoption and Advocacy. *Information Technology and Libraries*, 41(2), 1–17. <https://doi.org/10.6017/ital.v41i2.14683>
- Rolin, K., Kakkuri-Knuuttila, M.-L., Henttonen, E., Eräranta, K., & Gaudeamus oy (Toim.). (2006). *Soveltava yhteiskuntatiede ja filosofia*. Gaudeamus.
- Russell, S. J., Norvig, P., Chang, M.-W., Devlin, J., Dragan, A., Forsyth, D., Goodfellow, I., Malik, J., Mansinghka, V., Pearl, J., & Wooldridge, M. (2022). *Artificial intelligence: A modern approach* (Fourth edition). Pearson.
- Ryan, M., & Stahl, B. C. (2021). Artificial intelligence ethics guidelines for developers and users: Clarifying their content and normative implications. *Journal of Information, Communication & Ethics in Society (Online)*, 19(1), 61–86. <https://doi.org/10.1108/JICES-12-2019-0138>
- Salih, A. M., Galazzo, I. B., Raisi-Estabragh, Z., Petersen, S. E., Menegaz, G., & Radeva, P. (2024). Characterizing the Contribution of Dependent Features in XAI Methods. *IEEE Journal of Biomedical and Health Informatics*, 1–8. IEEE Journal of Biomedical and Health Informatics. <https://doi.org/10.1109/JBHI.2024.3395289>
- Schmidt, V., & Wood, M. (2019). Conceptualizing throughput legitimacy: Procedural mechanisms of accountability, transparency, inclusiveness and openness in EU governance. *Public Administration (London)*, 97(4), 727–740. <https://doi.org/10.1111/padm.12615>
- Schoonderwoerd, T. A. J., Jorritsma, W., Neerincx, M., A., & van den Bosch, K. (2021). Human-centered XAI: Developing design patterns for explanations of clinical

- decision support systems. *International Journal of Human-Computer Studies*, 154, 102684-. <https://doi.org/10.1016/j.ijhcs.2021.102684>
- Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), 1–31. <https://doi.org/10.1145/3419764>
- Shneiderman, B. (2022). *Human-Centered AI*. Oxford University Press USA - OSO.
- Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., & Elmqvist, N. (2018). *Designing the user interface: Strategies for effective human-computer interaction* (Sixth edition). Pearson.
- Stone, D. L., Deadrick, D. L., Lukaszewski, K. M., & Johnson, R. (2015). The influence of technology on the future of human resource management. *Human Resource Management Review*, 25(2), 216–231. <https://doi.org/10.1016/j.hrmr.2015.01.002>
- Tabrez, A., Agrawal, S., & Hayes, B. (2019). *Explanation-Based Reward Coaching to Improve Human Performance via Reinforcement Learning*. 2019-March, 249–257. Scopus. <https://doi.org/10.1109/HRI.2019.8673104>
- Tuomi, J., & Sarajärvi, A. (with Tammi (yhtiö)). (2018). *Laadullinen tutkimus ja sisällönanalyysi* (Uudistettu laitos). Kustannusosakeyhtiö Tammi.
- van Esch, P., Black, J. S., & Ferolie, J. (2019). Marketing AI recruitment: The next phase in job application and selection. *Computers in Human Behavior*, 90, 215–222. <https://doi.org/10.1016/j.chb.2018.09.009>

- Van Iddekinge, C. H., Lievens, F., & Sackett, P. R. (2023). Personnel selection: A review of ways to maximize validity, diversity, and the applicant experience. *Personnel Psychology, 76*(2), 651–686. <https://doi.org/10.1111/peps.12578>
- Vilkka, H. (with Santalahti-kustannus & PS-kustannus). (2021). *Tutki ja kehitä* (5., päivitetty painos). PS-kustannus.
- Vilone, G., & Longo, L. (2021). Classification of Explainable Artificial Intelligence Methods through Their Output Formats. *Machine Learning and Knowledge Extraction, 3*(3), 615–661. <https://doi.org/10.3390/make3030032>
- Vold, K. (2024). Human-AI cognitive teaming: Using AI to support state-level decision making on the resort to force. *Australian Journal of International Affairs, 78*(2), 229–236. <https://doi.org/10.1080/10357718.2024.2327383>
- Wiberg, M., & Stolterman Bergqvist, E. (2023). Automation of interaction—Interaction design at the crossroads of user experience (UX) and artificial intelligence (AI). *Personal and Ubiquitous Computing, 27*(6), 2281–2290. <https://doi.org/10.1007/s00779-023-01779-0>
- Woods, S. A., Ahmed, S., Nikolaou, I., Costa, A. C., & Anderson, N. R. (2020). Personnel selection in the digital age: A review of validity and applicant reactions, and future research challenges. *European Journal of Work and Organizational Psychology, 29*(1), 64–77. <https://doi.org/10.1080/1359432X.2019.1681401>
- Zhang, C. (2022). Evaluation and analysis of human resource management mode and its talent screening factors based on decision tree algorithm. *The Journal of Supercomputing, 78*(13), 15681–15713. <https://doi.org/10.1007/s11227-022-04499-z>

Liitteet

Liite 1. Haastattelurunko

Haastattelurunko

Alustavat kysymykset

Voitko kertoa, minkälaisissa työhakutilanteissa olet ollut mukana?

Oletko koskaan ollut mukana rekrytointiprosessissa, jossa tekoälyä käytettiin valintojen tekemiseen? Jos olet, voitko kuvata kokemustasi?

Selitettävyyks tekoälypohjaisissa rekrytointijärjestelmissä

1. Millaisia selityksiä odotat tekoälypohjaisilta rekrytointijärjestelmiltä?

Tulosperusteiset selitykset:

Haluaisitko saada perusteluja siitä, miksi sinut valittiin tai hylättiin rekrytointiprosessissa, jossa käytettiin tekoälyä?

Millainen tieto olisi sinulle hyödyllistä tekoälyjärjestelmän käsittelemästä datasta rekrytointiprosessin aikana?

Prosessiperusteiset selitykset:

Kuinka tärkeää sinulle on saada selityksiä rekrytointiprosessin vaiheista?

Haluaisitko tietoa siitä, milloin tekoälyä käytetään ja milloin ihmispäätöksiä tehdään?

Haluaisitko saada tarkempaa tietoa tekoälyn teknisistä ominaisuuksista, kuten turvallisuudesta ja suorituskyvystä?

Millaisia tietoja toivoisit saavasi tekoälyjärjestelmän oikeudenmukaisuuden varmistamisesta, kuten esimerkiksi syrjinnän estämisestä ja tasapuolisesta kohtelusta?

Millaisia tietoja haluaisit saada tekoälyjärjestelmän vaikutuksista, kuten siitä, miten järjestelmän suunnittelussa ja toteutuksessa on huomioitu sen mahdolliset vaikutukset yksilöihin ja yhteiskuntaan?

Teeman lopettava kysymys:

Pidätkö kokonaisuutena tärkeämpänä tulosperusteisia selityksiä (eli päätöksen lopputuloksen perusteluita) vai prosessiperusteisia selityksiä (eli sitä, miten päätökseen on päädytty)?

Selitettävyyden esittämistavat ja vuorovaikutuksellisuus

2. Miten tekoälypohjaisten rekryointijärjestelmien tekemät päätökset ja käsittelemät tiedot tulisi esittää työnhakijoille ymmärrettävästi rekryointiprosessin aikana?

Yleinen selitettävyys ja tiedon esittäminen:

Millaisia selityksiä odotat tekoälypohjaisilta rekryointijärjestelmiltä päätöksenteon eri vaiheissa?

Minkä tyyppisiä selityksiä koet tärkeimpinä rekryointijärjestelmän päätöksistä? Esimerkiksi: yksityiskohtaiset analyysit vai lyhyet yhteenvedot?

Voisitko antaa esimerkin tilanteesta, jossa selitys tekoälyn tekemästä päätöksestä olisi auttanut sinua ymmärtämään hakuprosessin paremmin?

Haluaisitko, että tekoälyjärjestelmä antaisi sinulle selityksen siitä, miksi jokin tietty tieto hakemuksestasi oli tärkeä päätöksenteon kannalta?

Miten tärkeää sinulle on, että tekoälypohjaisen järjestelmän tuottamat selitykset ovat helposti ymmärrettävissä ilman teknistä taustatietoa?

Olisiko sinulle hyödyllistä, jos rekrytointijärjestelmä selittäisi, miten se vertaa hakemustasi muihin hakijoihin?

Vuorovaikutuksellisuus:

Miten haluaisit olla vuorovaikutuksessa tekoälypohjaisen rekrytointijärjestelmän kanssa rekrytointiprosessin aikana?

Toivoisitko, että sinulla olisi mahdollisuus pyytää lisäselvityksiä tai tarkennuksia tekoälyn tekemistä päätöksistä? Jos kyllä, mitä haluaisit erityisesti tietää?

Miten haluaisit, että automaation rooli rekrytointiprosessissa selitettäisiin sinulle? Esimerkiksi: yksinkertaiset tiedotteet vai tarkemmat kuvaukset prosessin kulusta?

Haluaisitko saada tietoa siitä, milloin ja missä vaiheessa tekoäly tekee päätöksiä hakemuksesi käsittelyssä, ja milloin päätöksen tekee ihminen?

Liite 2. Selitettävyyden ja vuorovaikutuksellisuuden konseptuaalinen kehittäminen



Aihepiirianalyysi

Ensimmäinen vaihe: kerää aiheeseen liittyvää tietoa

1

Tunnista keskeiset käsitteet

Esimerkiksi

Rekryointiprosessi ja sen vaiheet

Tekoälypohjaiset rekryointijärjestelmät

Ennakoiva analytiikka

Algoritminen päätöksenteko

Keskustelevat käyttöliittymät

Selittävä tekoäly

Selitystietoinen vuorovaikutussuunnittelu

XAI-menetelmät



Kokonaisvaltainen ymmärrys



2

Muodosta kontekstikuvauksia ja konsepteja



Tunnista rekryointiprosessin vaiheet (hakeminen, seulonta, arviointi, valinta) ja arvioi, minne selitettävyyden tarve voi kohdistua kussakin vaiheessa.



Tunnista tekoälypohjaisen rekryointijärjestelmän toimintaympäristö, käyttäjät, ominaisuudet ja toiminnallisuudet.



Tunnista selitystietoiseen vuorovaikutussuunnitteluun liittyviä periaatteita ja toimintamalleja.

Vaatimusten kerääminen ja analysointi

Toinen vaihe: määritä vaatimukset ja muodosta niistä relevantteja käyttötapauksia

1

Kerää ja määritä vaatimukset

Esimerkki

Määritellyt vaatimukset

Selitetävyysodotukset

Konkreettisten kriteerien avaaminen

Tulosten vertailutiedot

Palaute kehityksen tueksi

Vuorovaikutusodotukset

Reaaliaikainen ja automaattinen tiedottaminen

Mahdollisuus kysyä lisätietoja

Automaation ja ihmisten roolien selkeyttäminen

2

Muodosta käyttötapaukset

Esimerkki

Käyttötapaus A

Rekryointivaihe

Hakemuksen jättämisen jälkeen

Vaatus

Konkreettisten kriteerien avaaminen

Reaaliaikainen ja automaattinen tiedottaminen

Käyttötapaus B

Rekryointivaihe

Ansioluettelon seulonta

Vaatus

Tulosten vertailutiedot

Reaaliaikainen ja automaattinen tiedottaminen



Skenaariot

Multimodaalinen vuorovaikutussuunnittelu

Kolmas vaihe: muodosta prototyyppi tai konseptuaalinen malli käyttötapauksiin

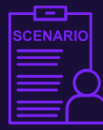
1 Suunnittele käyttötapauksille suunnittelumalli selityksille



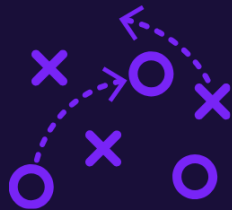
Tieto



Vaatimukset



Käyttötapaukset



Suunnittelumalli selityksille



Hyödynnä kerättyä tietoa, määriteltyjä vaatimuksia sekä erilaisia käyttötapauksia luodaksesi selitysmalleja, jotka tukevat rekrytointiprosessin eri vaiheita tekoälypohjaisessa järjestelmässä. Selitettävyyden voi kohdistaa hakemiseen, seulontaan, arviointiin ja valintaan.

2 Muodosta selittävä informaatio käyttäjälle testattavaksi



Muodosta prototyyppi tai konseptuaalinen malli, joka sisältää selittävää informaatiota käyttäjärhmälle. Testaa mallia käytännössä, jotta voit arvioida, miten hyvin selitykset tukevat ymmärrettävyyttä ja läpinäkyvyyttä rekrytointiprosessin eri vaiheissa.

Selitettävyyden kehityssuuntia



Konkreettisten valintakriteerien
avaaminen

Tulosten vertailutiedot



Palaute kehityksen tueksi

Prosessiperusteinen
läpinäkyvyys

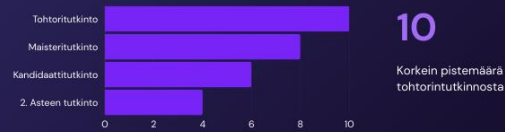


Konkreettisten valintakriteerien avaaminen

ESITÄ VALINTAKRITEERIEN PAINOTUKSIA



AVAA KRITEERIKOHTAISIA PISTEYTYKSIÄ



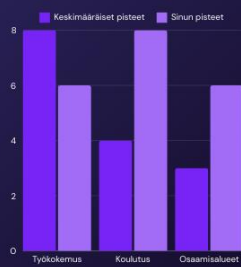
Pyri selittämään avoimesti, mitkä tekijät vaikuttivat rekryointipäätökseen tekoälypohjaisessa rekryointijärjestelmässä



Tunnista kriittisiä tietotyyppejä, kuten työkokemus, koulutustaso, osaamisalueet, ja pyri avaamaan niitä työnhakijoille ymmärrettävästi

Tulosten vertailutietoja

PISTEIDEN VERTAILU



TIETOA VERTAILUSTA

20
Sinun pisteet

25
Parhaat pisteet

14
Keskimääräinen pistemäärä

12–20.
Sijoitus hakemusten arvioinnin jälkeen

60%
Todennäköisyys haastatteluun kutsulle



Pyri tarjoamaan työnhakijoita kiinnostavaa tietopohjaa rekryointipäätöksiin liittyvistä tuloksista.



Esimerkiksi vertailutietojen tarjoaminen auttaa hakijaa ymmärtämään omaa sijoittumistaan kokonaisuuteen ja tunnistamaan mahdollisia kehityskohteita.

Palaute kehityksen tueksi

TIETOPOHJAINEN PALAUTE

| | |
|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| ● Koulutus | |
| ● Työkokemus toimialalla | Keskimääräisesti parempi kuin muilla hakijoilla |
| ● Työkokemus operatiivisen johtotason tehtävistä | |
| ● Puuttuva kielitaito: Espanja | Ansioluettelon seurantavaiheessa jäi huomaamatta merkityksellisiä valintakriteerejä näiltä osa-alueilta |
| ● Puuttuva osaamisalue: projektijohtaminen | |

PERUSTELEVA PALAUTE

Yleinen palaute

Saavutit 20 pistettä maksimipistemäärästä 30, mikä sijoittaa sinut hakijoiden joukossa välille 12 -20

Valinta

Kiitos hakemuksesta, mutta emme valitettavasti jatka rekrytointia kanssasi pidemmälle. Rekrytointi kohdallasi päättyi ansioluettelon automaattiseen seurantavaiheeseen.

Kehitysehdotukset

Koulutuksesi ja työkokemuksesi ovat hyvät pohjat alalle.

Voit parantaa mahdollisuuksiasi kehittämällä kielitaitoasi, erityisesti espanjan kielen osalta, sillä yrityksemme toimialue sijaitsee Espanjassa.



Pyri rakentamaan kehittävää ja informoivaa palautetta sisältävää viestintää

Prosessiperusteinen läpinäkyvyys



HYÖDYNNÄ XAI-MENETELMIÄ

XAI-menetelmät (Explainable AI) ovat tekoälyn selitettävyyden menetelmiä, joiden avulla voidaan ymmärtää ja selittää tekoälymallien päätöksiä ja toimintaa.

VAIKUTUKSET

Käytä XAI-menetelmiä (Esim. SHAP, LIME) selittämään, mitkä tekijät vaikuttavat hakijan arviointiin ja mahdollisiin vaikutuksiin

OIKEUDENMUKAISUUS

Esitä, miten XAI-menetelmien avulla pyritään, että tekoälymalli kohtelee eri ryhmiä tasapuolisesti ja estää syrjintää

VASTUULLISUUS

XAI-menetelmät tuo esiin päätöksenteon perusteet, mikä vahvistaa vastuullista tekoälyn käyttöä

TURVALLISUUS JA SUORITUSKYKY

Hyödynnä XAI-menetelmiä ennusteiden tarkasteluun, jotta turvallisuus- ja suorituskykyvaatimukset näkyvät päätöksenteossa

Vuorovaikutuksellisuuden kehityssuuntia

Tekoälypohjainen rekrytointi



Reaaliaikainen ja automaattinen tiedottaminen



Mahdollisuus kysyä lisätietoja

Automaation ja ihmisten roolien selkeyttäminen



Reaaliaikainen ja automaattinen tiedottaminen

ANSIOLUETTELON KONELUETTAVUUDEN ANALYSOINTI

Avainsanat



Työkokemus



Koulutustaso



Kielitaito: Järjestelmä ei tunnistanut kielitaitoon liittyviä tietoja



LinkedIn-linkki puuttuu

Ohjeet

Huomioithan, että dokumentin rakenne ja visuaaliset elementit voivat vaikuttaa koneluettavuuden analysointiin ja siten hakemusten seulontaan.

Täydennä kielitaito ja LinkedIn-linkki, jotta ansioluettelo olisi kattavampi



Tunnista, missä rekrytoinnin vaiheissa reaaliaikainen ja automaattinen tiedottaminen voi parhaiten tukea työnhakijan tavoitteita tekoälypohjaisissa rekrytointijärjestelmissä.



Pyri tarjoamaan rekrytoinnin kannalta merkityksellistä ja automaattisesti päivittyvää palautetta käsittelyn tueksi.

Mahdollisuus kysyä lisätietoja

KESKUSTELEVAT KÄYTTÖLIITTYMÄT



Voisitko avata tarkemmin hylkäykseen johtaneista päätöksistä?



Kiitos kysymyksestäsi! Hakemuksesi arvioinnissa käytettiin useita kriteerejä, kuten koulutustaso, työkokemus ja erityisosaamisalueet. Tässä tapauksessa päätös perustui siihen, että osa hakijoista täytti valintakriteerit hieman tarkemmin.

Haluatko lisätietoja vielä?

...



Työnhakijat arvostavat mahdollisuutta kysyä lisätietoja ja saada syvempiä tarkennuksia, erityisesti hylkäykseen johtaneissa rekrytointiprosesseissa



Keskustelevien käyttöliittymien konseptisuunnitteluun voidaan sisällyttää päätöksenteon selitettävyyden ja vuorovaikutukselliset elementit, kuten informaation välittäminen ja dialogi

Automaation ja ihmisten roolien selkeyttämien

VASTUUALUEIDEN SELKEYTTÄMINEN



Hakemuksen jättämien



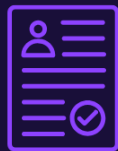
Tekoäly organisoii ja kommunikoi hakemiseen liittyvissä asioissa, kuten tarkistaa koneluettavuuden ja puuttuvia tietoja



Hakemuksen seulonta



Tekoäly suorittaa seulonnan määriteltyjen kandidaattikriteerien pohjalta



Hakemuksen arviointi



Rekrytoija arvioi seulontavaiheen läpäisseet hakemukset



Tekoälyä käytetään myös hakemusten arviointityökaluna



Haastattelu ja valinta



Rekrytoija suorittaa haastattelun ja lopullisen valinnan



Palautte



Tekoälyä käytetään palautteiden muodostamiseen seulontavaiheessa pudonneille



Loppuvaiheeseen suoriutuneet saavat myös palautetta rekrytoijilta



Esitä automaation ja ihmisen roolit selkeästi tekoälypohjaisessa rekryointiprosessissa- ja järjestelmässä



Voit esittää tarkemmin, miten automaatio käsittelee hakemuksia ja arvioi niitä

MATERIAALILINKKEJÄ SELITETTÄVYYDEN JA VUOROVAIKUTUKSELLISUUDEN SUUNNITTELUN KEHITTÄMISEKSI

[KONSEPTUAALINEN MALLI SELITETTÄVYYDEN SUUNNITTELEMISEKSI](#)

[SELITTÄVÄN TEKÖÄLYN VUOROVAIKUTUSSUUNNITTELU](#)

[TEKNINEN TIETO JA XAI-MENETELMÄT](#)