

Jere Valoma

# TEKOÄLYN SOVELTAMINEN HR-PROSESS- SIEN PÄÄTÖKSENTEOSSA

Keskeisiä eettisiä haasteita ja ratkaisuja

# TIIVISTELMÄ

Jere Valoma: Tekoälyn soveltaminen HR-prosessien päätöksenteossa: Keskeisiä eettisiä haasteita ja ratkaisuja  
Kandidaattitutkielma  
Tampereen yliopisto  
Tietojenkäsittelytieteiden tutkinto-ohjelma  
Maaliskuu 2025

---

Tässä kandidaattitutkielmassa tarkastellaan tekoälyn käytön eettisiä haasteita henkilöstöhallinnon (HR) prosesseissa, kuten rekrytoinnissa, työntekijäarvioinneissa ja koulutuksen suunnittelussa. Tutkimuksen tavoitteena on selvittää, millaisia riskejä ja ongelmia tekoälyn hyödyntäminen voi aiheuttaa HR-kontekstissa, erityisesti algoritmien puolueellisuuden, päätöksenteon läpinäkyvyyden sekä työntekijöiden yksityisyyden suojan näkökulmasta.

Kirjallisuuskatsauksen perusteella todetaan, että tekoälyjärjestelmät voivat vahvistaa rakenteellista syrjintää, jos niiden kehittämisessä käytetty data on puolueellista. Lisäksi ilmiö, jossa tekoälyn tekemät päätökset ovat vaikeasti ymmärrettäviä, heikentää työntekijöiden ja työnhakijoiden oikeusturvaa. Jatkuva digitaalinen seuranta puolestaan voi vaikuttaa negatiivisesti työhyvinvointiin. Tutkielmassa ehdotetaan useita keinoja ongelmien ratkaisemiseksi. Näihin kuuluvat muun muassa selitettävän tekoälyn (XAI) hyödyntäminen läpinäkyvyyden lisäämiseksi, monipuolisemman ja edustavamman koulutusdatan käyttö sekä hybridimallien kehittäminen, joissa ihmisten asiantuntemus säilyttää keskeisen roolinsa päätöksenteossa.

Vaikka tekoälyllä on merkittävä potentiaali tehostaa HR-prosesseja, sen käyttö edellyttää vastuullisia toimintamalleja ja asianmukaista sääntelyä. Tutkielma painottaa tarvetta jatkotutkimukselle, joka keskittyy eettisten periaatteiden konkreettiseen integrointiin ja tekoälyn vastuulliseen soveltamiseen organisaatioiden HR-järjestelmissä.

Avainsanat: tekoäly, tekoälyn eettiset haasteet, HR, rekrytointi, algoritmien puolueellisuus, yksityisyydensuoja

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

# TEKOÄLYN KÄYTTÖ OPINNÄYTTEESSÄ

Opinnäytteessäni on käytetty tekoälysovelluksia:

- Ei
- Kyllä

Ilmoitukseni mukaan olen käyttänyt opinnäytteessäni tutkielmaprosessin aikana seuraavia tekoälysovelluksia:

Tekoälysovellusten nimet ja versiot:

ChatGPT 4o

Käyttötarkoitus:

Tekoälyä hyödynnettiin rakenteen suunnittelussa. Kerroin mallille aiheideastani ja kandidaattitutkielman rajoitteista, jonka jälkeen pyysin sitä tekemään rungon tutkielmalleni. Sain tästä hyvän pohjan tutkielman rungolle, jota hienosäätämällä laadin lopullisen version.

Lisäksi hyödynsin tekoälyä kieliopin tarkastamiseen antamalla tekstipätkiäni sille ja pyytämällä palautetta kieliopista.

Osiot, joissa tekoälyä on käytetty:

Tutkielman rakenne

Kielioppi

Olen tietoinen siitä, että olen täysin vastuussa koko opinnäytteeni sisällöstä, mukaan lukien osat, joissa on hyödynnetty tekoälyä, ja hyväksyn vastuun mahdollisista eettisten ohjeiden rikkomuksista.

## SISÄLLYSLUETTELO

<b>1</b>	<b>Johdanto .....</b>	<b>1</b>
<b>2</b>	<b>HR-prosessit, tekoäly ja sen hyödyntäminen tietojärjestelmissä.....</b>	<b>2</b>
	2.1 HR-prosessien yleiskuvaus	2
	2.2 Tekoäly tietojärjestelmissä	3
<b>3</b>	<b>Tekoälyn nykytila ja sen potentiaali HR-järjestelmissä .....</b>	<b>4</b>
	3.1 Tekoälyllä saavutetut hyödyt HR-järjestelmissä	4
	3.2 Tekoälyn hyödyntämisen tulevaisuuden tavoitteet HR-järjestelmissä	5
<b>4</b>	<b>Keskeisiä eettisiä haasteita tekoälyn soveltamisessa HR-prosesseissa .....</b>	<b>6</b>
	4.1 Puolueellisuus ja syrjintä päätöksenteossa	6
	4.2 Läpinäkyvyyden puute ja epäinhimillisuus päätöksenteossa	7
	4.3 Yksityisyyden loukkaukset ja liiallisen valvonnan haitat	8
<b>5</b>	<b>Ratkaisuja eettisiin haasteisiin tekoälyn soveltamisessa HR-prosesseissa .....</b>	<b>10</b>
	5.1 Puolueellisuuden vähentäminen ja monimuotoisuudesta huolehtiminen	10
	5.2 Läpinäkyvyyden ja inhimillisyyden lisääminen tekoälypohjaisessa päätöksenteossa	12
	5.3 Yksityisyyden suojaaminen	14
<b>6</b>	<b>Johtopäätökset ja pohdinta .....</b>	<b>15</b>
	<b>Lähdeluettelo.....</b>	<b>17</b>

## 1 Johdanto

Tekoäly muuttaa työelämää ennennäkemättömällä vauhdilla. Yritykset hyödyntävät tekoälyä rekrytoinnista suorituskyvyn arviointiin ja työvoiman hallintaan, millä pyritään tehostamaan HR-prosesseja ja auttamaan organisaatioita tekemään parempia päätöksiä (Garg ym., 2021; Andrieux ym., 2024). Mutta mitä tapahtuu, kun nämä järjestelmät alkavat tehdä ratkaisuja, joita kukaan ei täysin ymmärrä? Voiko algoritmi olla puolueeton, vai heijastaako se vain menneisyyden vinoutumia? Entä mitä tapahtuu työntekijöiden yksityisyydelle, kun tekoäly analysoi heidän käyttäytymistään ja jopa tunnetilojaan?

Tutkimuksen päätavoitteena on sekä analysoida tekoälyn käytön eettisiä haasteita HR:n päätöksenteossa että esittää konkreettisia ratkaisuja näiden ongelmien hallintaan. Näiden tavoitteiden saavuttamiseksi tutkimus pyrkii vastaamaan tutkimuskysymyksiin, kuten millaisia eettisiä haasteita tekoälyn käyttö aiheuttaa HR-prosesseissa ja mitkä ovat tehokkaita keinoja vähentää tekoälyn käyttöön liittyviä eettisiä riskejä HR-prosesseissa. Kandidaattitutkielman rajallisuuden vuoksi työssä keskitytään tarkastelemaan kolmea keskeistä eettistä ongelmaa: algoritmien puolueellisuutta ja syrjintää, päätöksenteon läpinäkyvättömyyttä sekä työntekijöiden yksityisyydensuojaa. Aiheita tarkastellaan organisaation päätöksenteon ja HR-asiantuntijoiden näkökulmasta, eikä tutkielmassa siksi pu-reuduta teknologisiin yksityiskohtiin.

Tutkielman aineisto on kerätty systemaattisen tiedonhaun sekä alan asiantuntija-artikkeleista saadun taustatiedon avulla. Tieteellisen kirjallisuuden ja vertaisarvioitujen tutkimusjulkaisujen etsintään Google Scholar ja ScienceDirect (Elsevier) osoittautuivat tehokkaimmiksi työkaluiksi löytää relevanttia, uutta ja laadukasta tutkimustietoa. Lisäksi julkaisujen etsinnässä on hyödynnetty myös käänteistä lähdehakua (reverse literature search). Jotta kirjallisuuskatsaus olisi relevanttia ja luotettavaa, siihen on valittu vain vertaisarvioituja ja alan johtavissa julkaisukanavissa ilmestyneitä artikkeleita. Lisäksi tutkielmaan on valikoitu mukaan lähes ainoastaan 2020-luvulla julkaistuja lähteitä, sillä tekoälyn käyttö ja kehitys on edennyt viime vuosina huomattavan nopeasti. Näin on varmistettu, että teoriapohja ja tutkimukseen sisällytetty aineisto kuvastavat mahdollisimman hyvin nykytilannetta ja viimeisimpiä kehitysuuntia alalla.

## **2 HR-prosessit, tekoäly ja sen hyödyntäminen tietojärjestelmissä**

Tekoälyn hyödyntäminen HR-prosesseissa edellyttää ymmärrystä sekä HR:n perusprosesseista että teknologian roolista niiden tukena. Ennen kuin syvennymme tekoälyn vaihtuksiin ja eettisiin haasteisiin, esittelemme tiiviisti HR-prosessien kokonaisuutta ja sitä, miten tekoäly ilmenee tietojärjestelmissä.

### **2.1 HR-prosessien yleiskuvaus**

Goswamin (2018) mukaan nykypäivän HR:n prosessit koostuvat joukosta toisiinsa liittyviä toimintoja, jotka tukevat organisaation kykyä houkuttaa, kehittää, motivoida ja säilyttää työntekijöitä. Keskeisiä HR-prosesseja ovat rekrytointi ja valinta, koulutus ja kehitys, kyvykkyyksien hallinta, palkitseminen, työvoiman suunnittelu sekä työsuhdeasiat ja riskienhallinta. Rekrytointi ja valinta kattavat uusien työntekijöiden houkuttelun, seulonnan ja palkkaamisen. Koulutus ja kehitys liittyvät työntekijöiden taitojen parantamiseen ja heidän valmisteleminen tuleviin rooleihin. Kyvykkyyksien hallinta keskittyy kompetenssienarviointiin, tavoitteiden asettamiseen ja seurantaan. Palkitseminen sisältää sekä rahalliset että muut kannustimet työntekijöiden motivoimiseksi, kun taas työvoiman suunnittelu varmistaa oikeiden resurssien saatavuuden oikeaan aikaan. Työsuhdeasiat ja riskienhallinta keskittyvät työntekijöiden hyvinvointiin, turvallisuuteen ja organisaation sääntöjen noudattamiseen.

HR:llä on keskeinen rooli yrityksen menestyksen tukemisessa. Se ei ainoastaan hallinnoi organisaation tärkeintä resurssia eli ihmisiä, vaan myös varmistaa, että nämä ovat linjassa organisaation strategisten tavoitteiden kanssa. Hyvin toimiva HR edistää yrityksen kulttuuria, joka houkuttelee ja sitouttaa lahjakkaita työntekijöitä. Lisäksi HR-prosessit, kuten kyvykkyyksien hallinta ja koulutusohjelmat, varmistavat, että työntekijöiden taidot ja osaaminen kehittyvät vastaamaan yrityksen jatkuvasti muuttuvia tarpeita. (Goswami, 2018)

HR-tietojärjestelmillä (HRIS, human resources information system) on kriittinen rooli HR-prosessien tukemisessa ja tehostamisessa. HRIS mahdollistaa tietojen tehokkaan keräämisen, tallentamisen ja analysoinnin, mikä vähentää manuaalista työtä ja minimoi virheitä (Silva ym., 2017). Silva ym. (2017) korostavat, että HRIS integroi HR:n ja teknologian, mahdollistaen päätöksenteon, joka perustuu ajantasaiseen ja tarkkaan tietoon. HRIS tukee esimerkiksi strategista HR:ää auttamalla organisaatioita suunnittelemaan työvoimatarpeitaan ja tunnistamaan kehitysmahdollisuuksia. Järjestelmä myös

edistää työntekijöiden osallistumista esimerkiksi itsepalveluportaalien kautta, mikä on omiaan parantamaan työntekijöiden tyytyväisyyttä. (Silva ym., 2017)

## **2.2 Tekoäly tietojärjestelmissä**

Ongsuleen (2017) mukaan tekoäly on teknologia, joka jäljittelee ihmisen älykkyyttä analyysoimalla suuria tietomääriä, tunnistamalla kaavoja ja tekemällä itsenäisiä päätöksiä. Tekoälyn keskeisiä osa-alueita ovat koneoppiminen ja syväoppiminen. Koneoppiminen mahdollistaa järjestelmille oppimisen historiallisista tiedoista ilman ennalta määritettyjä sääntöjä, kun taas syväoppiminen hyödyntää neuroverkkoja monimutkaisten datarakenteiden analysointiin ja päätöksentekoon (Ongsulee, 2017; Sehwat, 2023). Tekoälyn keskeinen ominaisuus on sen kyky käyttää suuria määriä dataa koulutuksessa, jolloin algoritmit voivat jatkuvasti parantaa suorituskykyään ja tarkkuuttaan (Ongsulee, 2017; Sehwat, 2023).

Tekoälyä hyödynnetään laajasti erilaisissa tietojärjestelmissä. Esimerkkinä voidaan mainita yritysresurssien hallintajärjestelmät (ERP, enterprise resource planning), jotka keskittävät organisaation liiketoimintaprosessien hallinnan yhteen järjestelmään. Tekoälyn integrointi ERP-järjestelmiin mahdollistaa data-analyysin reaaliajassa, trendejä ennustavan analytiikan ja strategista päätöksentekoa tukevat toiminnallisuudet. Tekoäly voi esimerkiksi analysoida asiakasdatan pohjalta käyttäytymismalleja, jotka auttavat parantamaan asiakaskokemusta ja optimoimaan toimitusketjun hallintaa. (Bawa, 2023; Sehwat, 2023)

Keskeinen ero tekoälyn ja perinteisen automaation välillä on Sehwatin (2023) mukaan tekoälyn kyky oppia ja mukautua ympäristöönsä. Perinteinen automaatio suorittaa ennalta ohjelmoituja tehtäviä tehokkaasti, mutta se ei kykene reagoimaan muuttuviin tilanteisiin ilman ihmisen väliintuloa. Tekoäly puolestaan analysoi uusia tietoja, tekee niiden perusteella johtopäätöksiä ja sopeutuu tilanteen mukaan, mikä tekee siitä erityisen hyödyllisen monimutkaisissa ja dynaamisissa järjestelmissä. (Ongsulee, 2017; Sehwat, 2023)

Tekoälyn käyttö tietojärjestelmissä, kuten ERP:ssä, osoittaa sen potentiaalin muuttaa järjestelmät pelkistä operatiivisista työkaluista älykkäiksi ratkaisuiksi. Tämä mahdollistaa resurssien paremman hyödyntämisen, prosessien optimoinnin ja päätöksenteon tarkkuuden lisäämisen, mikä tekee tekoälystä keskeisen teknologian tietojärjestelmien kehityksessä. (Bawa, 2023; Sehwat, 2023)

### **3 Tekoölyn nykytila ja sen potentiaali HR-järjestelmissä**

Tekoölyn käyttö HR-järjestelmissä on viime vuosina lisääntynyt. Ennen kuin perehdymme tekoölyn käytön eettisiin haasteisiin, on tärkeää ymmärtää tekoölyn nykytilanne HR-kontekstissa ja millaisia hyötyjä sen soveltaminen on tähän mennessä tuonut. Tämä kappale auttaa hahmottamaan, millaista hyötyä tekoölyllä voidaan saada ja millaisia kehityssuuntia sille ennakoidaan tulevaisuudessa.

#### **3.1 Tekoölyllä saavutetut hyödyt HR-järjestelmissä**

Tekoöly on yhä keskeisemmässä roolissa HR-järjestelmien kehittämisessä ja käytössä (Garg ym., 2021). Sudheer Devarajun (2024) artikkeli kokoaa yhteen kattavasti aiempaa tutkimustietoa tekoölypohjaisten HR-järjestelmien vaikutuksista julkishallinnossa, esittäen yksityiskohtaisia tutkimustuloksia.

HR-järjestelmissä tekoölyä hyödynnetään laajasti rekrytoinnissa ja valinnassa. Ansioluetteloiden analysoinnissa ja ehdokkaiden esikarsinnassa tekoöly on lyhentänyt käsittelyaikaa jopa 92,3 % ja parantanut ehdokkaiden laadun arviointia 95,2 %:n tarkkuuteen. Lisäksi tekoölyn käyttö rekrytointiviestinnässä auttaa tunnistamaan ja vähentämään syrjiviä kielellisiä malleja, mikä on pienentänyt syrjintätapausten määrää 67,2 %. (Devaraju, 2024; Garg ym., 2021)

Työvoiman suunnittelu ja analytiikka on toinen merkittävä osa-alue, jossa tekoölyä käytetään ennakoimaan henkilöstötarpeita ja optimoimaan resurssien käyttöä. Tekoölyjärjestelmien avulla voidaan käsitellä laajoja tietomassoja, kuten eläköitymisennusteita ja taitomatriiseja, ja ennustaa henkilöstöresurssien tarve jopa 93 %:n tarkkuudella kahdeksan kuukauden aikajänteellä. (Devaraju, 2024; Garg ym., 2021)

Suorituskyvyn hallinta hyötyy tekoölyn kyvystä käsitellä työntekijöiden tuottavuusdataa. Syväoppimista hyödyntävät järjestelmät parantavat työntekijöiden suorituskyvyn ennustettavuutta 37,2 % ja tunnistavat epätavalliset suorituskykytrendit 97,1 %:n tarkkuudella, mikä edistää parempaa päätöksentekoa esimerkiksi koulutustarpeiden määrittämisessä. (Devaraju, 2024; Garg ym., 2021)

Työntekijäkokemuksen parantaminen on myös tekoölyn hyödyntämisen keskiössä. Luonnollisen kielen prosessointia (NLP, natural language processing) hyödyntämällä voidaan analysoida työntekijöiden palautetta ja tunnistaa avaintekijöitä, jotka vaikuttavat tyytyväisyyteen. Tämä on lisännyt työntekijöiden sitoutumisen mittauksen tarkkuutta lähes 55 %. (Devaraju, 2024; Garg ym., 2021)

Devarajun (2024) tutkimuksen pohjalta voidaan todeta, että tekoälyllä on potentiaali tehostaa merkittävästi HR-prosessien eri osa-alueita. Se voi nopeuttaa päätöksentekoa, parantaa arviointien tarkkuutta ja mahdollistaa työntekijäkokemuksen syvällisemmän analysoinnin.

### **3.2 Tekoälyn hyödyntämisen tulevaisuuden tavoitteet HR-järjestelmissä**

Tekoälyn soveltamisella HR-järjestelmiin tavoitellaan jatkossakin lisähyötyjä, joita ovat Nyathanin (2023) ja Okattan ym. (2024) mukaan muun muassa HR-prosessien ennakoivuuden parantaminen, työntekijäkokemusten yksilöllistäminen sekä teknologian ja eettisten käytäntöjen synergian vahvistaminen.

Tekoäly voi auttaa HR-toimintoa kehittämään kykyään ennakoida organisaation tulevia tarpeita. Ennakoivan analytiikan avulla voidaan esimerkiksi arvioida taitovaatimuksia ja työntekijöiden käyttäytymismalleja, mikä auttaa suunnittelemaan tulevaisuuden työvoimastrategioita. Tämä antaa organisaatioille mahdollisuuden sopeutua nopeasti muuttuvaan toimintaympäristöön ja varmistaa kilpailukykyä. (Nyathani, 2023; Okatta ym, 2024)

Tekoälyn edistyneet ominaisuudet, kuten luonnollisen kielen käsittely ja sentimenttianalyysi, voivat merkittävästi parantaa työntekijäkokemusta tarjoamalla yksilöllisiä ratkaisuja ja sitouttamisstrategioita (Nyathani, 2023). Esimerkiksi NLP:n avulla voidaan analysoida työntekijöiden palautetta reaaliajassa, minkä ansiosta on mahdollista tunnistaa huolenaiheet ja kohdentaa toimenpiteitä entistä tarkemmin (Nyathani, 2023). Lisäksi tekoälyn avulla voidaan räätälöidä työntekijöiden koulutusohjelmia ja urapolkuja heidän yksilöllisten tarpeidensa sekä vahvuuksiensa perusteella, mikä lisää sekä motivaatiota että tuottavuutta (Okatta ym., 2024).

Nyathani (2023) kertoo, että tekoäly myös tukee jatkuvaa innovointia ja iteratiivista kehitystä HR-strategioissa. Reaaliaikaiset palautesilmukat, joita tekoäly mahdollistaa, auttavat organisaatioita reagoimaan nopeasti henkilöstöön liittyviin muutoksiin ja parantamaan HR-käytäntöjään jatkuvasti. Tämä tekee organisaatioista ketterämpiä ja paremmin tulevaisuuden haasteisiin valmistautuneita.

Kuitenkin Okattan ym. (2024) ja Nyathanin (2023) molempien artikkeleissa korostetaan tekoälyn ja eettisten käytäntöjen välistä synergiaa HR-järjestelmien kehittämisessä. Tekoälyn laajamittainen käyttöönotto HR:ssä vaatii eettisiä hallintamekanismeja, kuten vaikutusten jatkuvaa arviointia ja työntekijöiden osallistamista. Näiden käytäntöjen tarkoituksena on vahvistaa työntekijöiden luottamusta järjestelmiin sekä varmistaa, että

tekoälyä hyödynnetään vastuullisesti ja oikeudenmukaisesti, mikä puolestaan tukee HR:n pitkäjänteistä kestävyyttä. (Nyathani, 2023; Okatta ym, 2024)

## **4 Keskeisiä eettisiä haasteita tekoälyn soveltamisessa HR-prosesseissa**

Vaikka tekoäly voi parantaa HR-prosessien tehokkuutta ja päätöksenteon laatua, sen käyttö voi tuoda mukanaan myös merkittäviä eettisiä haasteita. Seuraavaksi esitellään keskeisiä eettisiä ongelmia, jotka liittyvät tekoälyn käyttöön HR:ssä, ja pohditaan ongelmien taustalla vaikuttavia tekijöitä.

### **4.1 Puolueellisuus ja syrjintä päätöksenteossa**

Rigotti & Fosch Villarongan (2024) mukaan yksi merkittävimmistä huolenaiheista käytönotossa on tekoälyjärjestelmien taipumus puolueellisuuteen ja syrjintään esimerkiksi sukupuolen, iän tai etnisen taustan perusteella. Tekoälyjärjestelmien puolueellisuus ei ole satunnainen ilmiö, vaan se juontaa juurensa tiettyihin rakenteellisiin ja teknisiin tekijöihin, jotka vaikuttavat algoritmien toimintaan. Bar-Gil ym. (2024), Rigotti & Fosch-Villaronga (2024) ja Andrieux ym. (2024) jakavat yhteisen käsityksen siitä, että puolueellisuus syntyy tyypillisesti kolmella erilaisella tavalla: historiallisen datan vinoumista, tietojoukkojen representaatio-ongelmista sekä algoritmien ja mallien suunnittelusta. Näiden tekijöiden ymmärtäminen on keskeistä, jotta voidaan arvioida, miksi ja miten tekoäly saattaa johtaa syrjiviin päätöksiin HR-prosesseissa.

Andrieux ym. (2024) esittää, että tekoälymallit oppivat historiallisen datan pohjalta, ja jos tämä data sisältää syrjiviä käytäntöjä tai rakenteellisia vinoumia, tekoäly saattaa toistaa ja jopa vahvistaa niitä. Esimerkiksi, jos yrityksen aiemmissa rekrytoinneissa on suosittu tiettyjä hakijaryhmiä, kuten miesten yliedustus teknisissä tehtävissä, voi tekoälyjärjestelmä oppia tämän mallin ja alkaa automaattisesti suosia miehiä jatkossakin (Andrieux ym., 2024; Rigotti & Fosch-Villaronga, 2024). Tunnettu esimerkki tästä on Bar-Gilin ym. (2024) esittelemä maailmanlaajuisen suuryrityksen Amazonin rekrytointialgoritmi, joka analysoi yrityksen aiempia rekrytointipäätöksiä ja huomasi, että teknisiin rooleihin oli palkattu enimmäkseen miehiä. Tämän seurauksena järjestelmä alkoi automaattisesti alentaa hakijoiden pisteytystä, jos ansioluettelossa esiintyi esimerkiksi sana ”women’s” (esim. ”women’s chess club captain”) tai jos hakija oli valmistunut naisille suunnatusta yliopistosta. Vaikka Amazon lopulta luopui algoritmista, tapaus havainnollistaa,

kuinka tekoälyn päätöksenteko voi perustua menneisiin vinoumiin ja johtaa syrjiviin käytäntöihin.

Algoritminen syrjintä ei kuitenkaan aina johdu pelkästään historiallisen datan vinoumista (Andrieux ym., 2024). Tekoälyn oppimisprosessi on nimittäin hyvin riippuvainen koulutusdatan kattavuudesta ja tasapuolisuudesta (Rigotti & Fosch-Villaronga, 2024). Tämä riippuvuus saattaa johtaa tietojoukkojen representaatio-ongelmaan, jossa tietyt ryhmät ovat aliedustettuina datassa, ja minkä takia tekoäly voi tehdä niihin liittyviä päätöksiä syrjivällä tavalla (Andrieux ym., 2024). Lisäksi tämä voi johtaa esimerkiksi siihen, että vähemmistöt eivät tule valituiksi työhaastatteluihin, koska tekoäly ei osaa tunnistaa heidän potentiaaliaan yhtä hyvin kuin enemmistöryhmien hakijoiden kohdalla (Rigotti & Fosch-Villaronga, 2024). Erityisen ongelmallisia ovat kasvojentunnistukseen perustuvat rekrytointijärjestelmät, jossa monet tekoälymallit tunnistavat valkoihoisten miesten kasvot huomattavasti paremmin kuin muiden etnisten ryhmien edustajien kasvot. Tämä on seurausta siitä, että useimmat tekoälymallit koulutetaan pääosin länsimaisten mieshakijoiden kuvilla, mikä johtaa heikompaan tunnistustarkkuuteen muille ryhmille. (Bar-Gil ym., 2024)

Vaikka data on keskeinen tekijä tekoälyn puolueellisuuden muodostumisessa, myös algoritmien suunnitteluun liittyvät valinnat voivat vaikuttaa lopputuloksiin. Algoritmit voivat vahvistaa ennakkoluuloja, jos ne perustuvat harhaanjohtaviin attribuutteihin, kuten hakijan nimeen tai asuinpaikkaan (Rigotti & Fosch-Villaronga, 2024). Esimerkiksi jos rekrytointialgoritmi oppii, että tietystä kaupunginosasta tulevat hakijat ovat historiallisesti menestyneet heikommin tietyissä tehtävissä, järjestelmä voi alkaa suoraan aliarvioida näiden hakijoiden potentiaalia ilman objektiivista perustetta (Andrieux ym., 2024).

Lisäksi tekoälyjärjestelmien kehittäjät saattavat optimoida tekoälyn päätöksentekoa väärin mittareiden perusteella. Jos esimerkiksi rekrytointialgoritmi keskittyy vain aiempaan suorituskyydataan eikä huomioi hakijoiden potentiaalia tai oppimiskykyä, se voi estää monimuotoisuuden edistämistä ja syrjiä uransa alkuvaiheessa olevia hakijoita. (Bar-Gil ym., 2024)

## **4.2 Läpinäkyvyyden puute ja epäinhimillisyys päätöksenteossa**

Tekoälyn käyttö HR:n päätöksenteossa on herättänyt huolta erityisesti läpinäkyvyyden ja inhimillisen kontaktin puutteen osalta, sillä tekoälyn päätöksentekoprosessit ovat usein niin monimutkaisia, että työntekijöiden on vaikea ymmärtää millä perusteilla esi-

merkiksi rekrytoinnit, suoritusarvioinnit tai ylennykset tehdään (Bankins ym., 2022; Rodgers ym., 2023). Tätä ilmiötä, jota on kutsuttu mustan laatikon ongelmaksi (Hassija ym., 2024). Mustan laatikon ongelma korostuu erityisesti tekoälyn tehdessä työntekijöiden kannalta merkittäviä ratkaisuja, kuten irtisanomisia tai ylennyksiä, ilman selkeää perustelua (Hassija ym., 2024; Rodgers ym., 2023). Jos tekoäly ei tarjoa selkeitä perusteluja päätöksilleen, työntekijät saattavat kokea päätökset sattumanvaraisina ja epäoikeudenmukaisina. Tämä heikentää työntekijän luottamusta organisaatioon ja lisätä epävarmuuden tunnetta. (Bankins ym., 2022; Rodgers ym., 2023)

Työntekijän kokema epäselvyys päätöksistä voi vaikuttaa negatiivisesti organisaatiokulttuuriin. Päätöksenteon perusteet voivat herättää epäilyksiä mahdollisesta puolueellisuudesta, vaikka algoritmit olisivat teknisesti neutraaleja. Tästä voi syntyä epävarmuutta ja dehumanisaation tunne, jossa työntekijä kokee tulevansa mitätöidyksi pelkäksi datapisteeksi. Tämän kaltaiset tunteet voivat vähentää työtyytyväisyyttä ja sitoutumista sekä voivat jopa aiheuttaa psykologista stressiä. (Bankins ym., 2022; Rodgers ym., 2023)

Perinteisesti HR-prosessien vahvuutena on ollut mahdollisuus keskustella päätöksistä esihenkilöiden ja HR-asiantuntijoiden kanssa (Bankins ym., 2022). Tekoälyä käytettäessä tämä inhimillinen vuorovaikutus katoaa, mikä vaikeuttaa työntekijöiden mahdollisuutta saada henkilökohtaista palautetta ja tukea esimerkiksi negatiivisten päätösten, kuten hylättyjen ylennysanomusten tai kielteisten suoritusarviointien yhteydessä (Bankins ym., 2022; Mehrotra, 2024). Tällainen päätöksenteko, joka ei huomioi yksilöllisiä olosuhteita, voi tuntua epäinhimilliseltä ja johtaa kokemukseen siitä, ettei työntekijän panosta arvosteta.

Yhteenvedona voidaan todeta, että tekoälypohjainen päätöksenteko HR:ssä herättää huolta läpinäkyvyyden ja inhimillisen kontaktin puutteesta, joka saattaa heikentää työntekijöiden motivaatiota, sitoutumista ja luottamusta organisaatioon. Edellä kuvattu korostaa tarvetta kehittää järjestelmiä, jotka yhdistävät teknologian tehokkuuden ja inhimillisen harkinnan, jotta päätöksenteon oikeudenmukaisuus ja yksilöllisyys säilyvät.

### **4.3 Yksityisyyden loukkaukset ja liiallisen valvonnan haitat**

Tekoälyjärjestelmät keräävät ja analysoivat työntekijöistä laajoja tietomääriä, kuten suoritustietoja, viestintämalleja ja jopa tunnetiloihin liittyvää dataa. Tämä voi haastaa yksityisyyden suojan, erityisesti kun kerätyistä tiedoista ja niiden käyttötarkoituksista ei välttämättä aina tiedoteta. (Roemmich ym., 2023; Andrieux ym., 2024)

Andrieux ym. (2024) kertoo, että yksi suurimmista tietosuojariskeistä liittyy tekoälyjärjestelmien kykyyn käsitellä ja yhdistellä laajoja tietomääriä. Nämä järjestelmät voivat tuottaa johtopäätöksiä työntekijöistä ja heidän soveltuvuudestaan erilaisiin työtehtäviin analysoimalla heidän käyttäytymistään ja vuorovaikutustaan. Tämä voi johtaa tilanteisiin, joissa työntekijöistä tehdään päätelmiä ilman heidän suostumustaan ja ilman, että heillä on mahdollisuutta vaikuttaa tehtyjen arviointien perusteisiin (Andrieux ym., 2024). Vaikka Bar-Gilin ym. (2024) mukaan tekoälyn avulla voidaan saavuttaa tarkempia ja objektiivisempia arvioita, voi tietosuoja ongelmien ratkaisematta jättäminen lisätä epäluotamusta työntekijöiden ja organisaation välillä. Mikäli työntekijät kokevat, että heidän yksityisyyttään loukataan tai että heidän henkilötietojaan käytetään ilman asianmukaista suojaa, saattavat he alkaa rajoittaa omaa toimintaansa työpaikalla tavalla, joka ei ole organisaation tai heidän itsensä kannalta optimaalista.

Tekoälyn avulla ei pyritä pelkästään tehostamaan HR-prosesseja, vaan sillä pyritään myös luomaan uusia tapoja johtaa ja seurata työvoimaa. Roemmich ym. (2023) esittelevät emotion AI -teknologian, joka on tekoälyn sovellus, jonka tarkoituksena on analysoida ja tulkita ihmisten tunnetiloja erilaisten fysiologisen ja käyttäytymiseen liittyvän datan, kuten kasvonilmeiden, äänensävyn, puhutavan ja kehon kielen, perusteella. Tätä teknologiaa käytetään työpaikoilla esimerkiksi työntekijöiden stressitasojen, motivaation ja työtyytyväisyyden arvioimiseen.

Reaaliaikainen tekoälyanalyysi lisää työnantajan kontrollia, mutta samalla se voi heikentää työntekijöiden autonomiaa (Roemmich ym., 2023). Perinteisesti suorituskykyä on arvioitu esihenkilöiden havainnoin ja keskustelujen avulla, mutta jatkuva monitorointi mahdollistaa jokaisen työtoiminnan analysoinnin ja pisteyttämisen, mikä voi aiheuttaa jatkuvan tarkkailtuna olon tunteen ja lisätä psykologista painetta (Roemmich ym., 2023; Bar-Gil ym., 2024). Erityisesti emotion AI -teknologiat herättävät huolta jatkuvan valvonnan vaikutuksista. Roemmich ym. (2023) esittävät, että nämä järjestelmät voivat analysoida työntekijöiden kasvonilmeitä, äänenpainoa ja puhetyyliä tunnistakseen heidän tunnetilojaan, vaikka työntekijällä itsellään ei olisi mahdollisuutta kontrolloida tai vaikuttaa analyysin lopputulokseen. Tällainen valvonta voi aiheuttaa työntekijöille stressiä ja epävarmuutta, koska heidän reaktioitaan voidaan tulkita virheellisesti tai niitä voidaan käyttää päätöksenteossa ilman, että työntekijät voivat itse selittää käyttäytymistään. Lisäksi jatkuva tekoälyvalvonta voi muuttaa työntekijöiden käyttäytymistä tavalla, joka ei ole luonnollista tai työn kannalta tehokkainta. Jos työntekijät tietävät, että heidän suoritustaan analysoidaan jatkuvasti tekoälyjärjestelmän toimesta, he saattavat mukauttaa

työskentelyään vastaamaan järjestelmän oletettuja vaatimuksia sen sijaan, että he tekisivät työnsä tehokkaimmalla tai luovimmalla mahdollisella tavalla. Tämä voi johtaa esimerkiksi riskinottohalukkuuden vähenemiseen ja organisaatiokulttuurin jäykistymiseen. (Bar-Gil ym., 2024; Roemmich ym., 2023)

Roemmichin ym. (2023) mukaan tekoälypohjaisen valvonnan lisääntyminen voi vaikuttaa merkittävästi työntekijöiden henkiseen hyvinvointiin. Jos työntekijöiden käyttäytymistä analysoidaan ja arvioidaan jatkuvasti, voi se lisätä työperäistä stressiä ja vaikuttaa negatiivisesti työtyytyväisyyteen. Erityisesti emotion AI -teknologian käyttö voi johtaa siihen, että työntekijät kokevat painetta esittää tietämykseen perustuvia tunteilmaisuuksia työpaikalla, vaikka se ei vastaisi heidän todellista tunnetilaansa. Tämä ilmiö, jota kutsutaan emotionaaliseksi pakkotyöksi, voi heikentää työntekijöiden hyvinvointia ja johtaa pitkäaikaisiin psykologisiin haittoihin. Lisäksi emotion AI -teknologian tuottamat analyysit voivat aiheuttaa epäoikeudenmukaisia odotuksia työntekijöille, jos analyysit käytetään mittareina suoritusarvioinneissa. Esimerkiksi järjestelmän tunnistessa työntekijän tunnetilan usein "väsyneeksi" tai "ärtyneeksi", voi tämä vaikuttaa työntekijän saamiin uramahdollisuuksiin tai suoritusarviointeihin negatiivisesti, vaikka hänen suoriutumisensa työtehtävissä olisi objektiivisesti hyvä. Työntekijä voi tämän seurauksena kokea turhautumista siitä, että häntä arvioidaan tekoälyn tekemän virheellisen tulkinnan perusteella.

## **5 Ratkaisuja eettisiin haasteisiin tekoälyn soveltamisessa HR-prosesseissa**

Luvussa 4 käsitellyt eettiset haasteet osoittavat, että tekoälyn käyttö HR-prosesseissa ei ole aina ongelmaton. Jotta sen hyödyt voidaan saavuttaa ilman merkittäviä haittoja, on tärkeää kehittää ja ottaa käyttöön ratkaisuja, jotka edistävät oikeudenmukaista ja läpinäkyvää tekoälyn soveltamista. Tässä luvussa tarkastelemme keinoja, joilla voidaan vähentää puolueellisuutta, lisätä läpinäkyvyyttä ja suojella työntekijöiden yksityisyyttä tekoälypohjaisessa päätöksenteossa.

### **5.1 Puolueellisuuden vähentäminen ja monimuotoisuudesta huolehtiminen**

Seuraavaksi keskitytään ratkaisuihin, joilla tekoälyn puolueellisuutta voidaan vähentää ja monimuotoisuutta edistää HR-prosesseissa. Esitettävien ratkaisujen kautta pyritään varmistamaan, että tekoäly toimii HR-prosesseissa tukena eikä esteenä tasa-arvoi-

selle ja monimuotoiselle työelämälle. Rodgers ym. (2023) korostavat, että tekoälyjärjestelmien oikeudenmukaisuutta voidaan parantaa yhdistämällä teknisiä ratkaisuja, eettisiä päätöksentekomalleja ja organisaatiotason linjauksia.

Andrieux ym. (2024) esittää, että yksi tärkeimmistä tavoista vähentää tekoälyn puolueellisuutta on varmistaa, että käytettävä data edustaa tasapuolisesti eri tarkasteltavia ryhmiä. Tietojoukkojen tulee olla riittävän laajoja ja kattavia, jotta tekoäly pystyy tekemään tarkkoja ja oikeudenmukaisia päätöksiä myös aliedustettujen ryhmien kohdalla. Esimerkiksi rekrytointialgoritmeissa koulutusdatan tulee sisältää hakemuksia monipuolista taustoista tulevilta henkilöiltä, jotta järjestelmä ei kehittä ennakkoluuloja esimerkiksi sukupuolen, iän tai etnisen taustan perusteella (Andrieux ym., 2024).

Bar-Gil ym. (2024) puolestaan korostavat, että tietojoukkojen representaatio-ongelmat voivat aiheuttaa rekrytointipäätösten vääristymistä, jos tekoäly ei ole koulutettu tunnistamaan eri taustoista tulevien hakijoiden potentiaalia. Siksi on tärkeää käyttää erilaisia datan tasapainotusmenetelmiä, joissa aliedustettuja ryhmiä edustavan datan määrää kasvatetaan tasapuolisuuden varmistamiseksi.

Teknisten ratkaisujen osalta algoritmien vinouma-analyysi nousee keskiöön. Rigotti ja Fosch-Villaronga (2024) ehdottavat säännöllisiä auditointeja, joissa tekoälyn päätöksiä testataan eri ryhmien kohdalla. Tämä tekninen toimenpide voidaan nähdä täydentävänä Rodgers ym. (2023) esittämille eettisille päätöksentekomalleille, joissa päätöksenteon tulisi perustua paitsi tehokkuuteen myös sosiaalisten ja eettisten vaikutusten arviointiin. Yhdistämällä auditointimenetelmät ja eettiset mallit saadaan kokonaisvaltaisempi lähestyminen tekoälyn puolueellisuuden ehkäisyyn.

Rodgersin ym. (2023) esittämien eettisten päätöksentekomallien lisäksi voidaan myös hyödyntää käytännönläheisiä menetelmiä eettisen päätöksenteon tueksi. Esimerkiksi tästä voi nostaa Andrieuxin ym. (2024) Two-Rule Methodin. Tämä menetelmä perustuu kahteen periaatteeseen: “Do No Harm” (Älä aiheuta haittaa) ja “Do Good” (Tee hyvää). Ensimmäinen sääntö edellyttää, että tekoälyn käyttö ei saa vahingoittaa työntekijöitä tai työnhakijoita, esimerkiksi luomalla syrjiviä rekrytointikäytäntöjä tai vahvistamalla organisatorisia vinoumia. Toinen sääntö korostaa, että tekoälyä tulisi hyödyntää vain silloin, kun sen käyttö tuo enemmän hyötyä kuin haittaa, esimerkiksi tehostamalla rekrytointiprosessia ilman, että oikeudenmukaisuus ja läpinäkyvyys vaarantuvat (Andrieux ym., 2024). Two-Rule Methodin yksinkertaisuus tekee siitä käyttökelpoisen työkalun organisaatioille, jotka pyrkivät vastuulliseen tekoälyn hyödyntämiseen HR:ssä. Se

tarjoaa selkeän ja helposti omaksuttavan lähestymistavan, joka voi auttaa yrityksiä tasa-painottamaan teknologian tehokkuuden sekä eettiset vaatimukset keskenään.

Rodgers ym. (2023) mukaan tekoälypohjaisen päätöksenteon tulisi myös olla linjassa organisaation kokonaisstrategian kanssa. Rodgers ym. (2023) painottavat, että yritysten on asetettava monimuotoisuuden ja tasa-arvon edistäminen tekoälyjärjestelmiensä tavoitteeksi samalla kun ne huomioivat liiketoiminnalliset tarpeensa. Tällöin tekoäly ei optimoi päätöksiä pelkästään tehokkuuden näkökulmasta, vaan se huomioi myös HR:n eettiset ulottuvuudet.

Puolueellisuuden ehkäisy ei ole kertaluonteinen prosessi, vaan vaatii jatkuvaa seuranta- ja kehittämistä. Säännölliset algoritmien auditoinnit ovat tärkeä keino varmistaa, että tekoälyjärjestelmät eivät ala ajan myötä kehittää uusia vinoumia (Rodgers ym., 2023; Andrieux ym., 2024). Tämä tarkoittaa, että yritysten on otettava käyttöön kontrollimekanismeja, joilla analysoidaan rekrytointialgoritmien päätöksiä ja vertaillaan niitä manuaaliseen arviointiin (Rodgers ym., 2023). Lisäksi Andrieux ym. (2024) sekä Hassija ym. (2024) korostavat, että HR-ammattilaisten on osallistuttava aktiivisesti tekoälyn päätöksentekoprosessien arviointiin, jotta mahdolliset puolueellisuudet havaitaan varhaisessa vaiheessa. Eräs keino heidän mukaansa on käyttää human-in-the-loop -mallia, jossa tekoälyn tekemät alustavat päätökset tarkistetaan ihmisten toimesta ennen lopullista hyväksyntää. Tämä vähentää tekoälyn autonomisuutta, mutta mahdollistaa tarvittaessa korjaavat toimenpiteet ennen päätösten tekemistä.

## **5.2 Läpinäkyvyyden ja inhimillisyyden lisääminen tekoälypohjaisessa päätöksenteossa**

Tekoälypohjaisessa päätöksenteossa keskeisiä ongelmia ovat läpinäkyvyyden puute ja epäinhimillisyyteen liittyvät ongelmat, kuten algoritmien mustan laatikon ongelma, vuorovaikutuksen puute ja työntekijöiden kokema dehumanisaatio (Bankins ym., 2022; Hassija ym., 2024; Mehrotra, 2024; Rodgers ym., 2023). Tässä kappaleessa käsitellään keinoja, joilla tekoälypohjaisen päätöksenteon läpinäkyvyyttä voidaan lisätä, työntekijöiden ja HR-ammattilaisten luottamusta vahvistaa sekä varmistaa, että tekoälyjärjestelmät tukevat oikeudenmukaista ja inhimillistä työelämää.

Tekoälypohjaisten päätösten tulisi olla selitettäviä ja jäljitettävissä, jotta työntekijät ja HR-ammattilaiset voivat ymmärtää ja arvioida niiden oikeudenmukaisuutta. Selitettävä tekoäly (Explainable AI, XAI) on yksi keskeinen ratkaisu tähän ongelmaan, sillä se mah-

dollistaa päätöksenteon logiikan avaamisen käyttäjille ymmärrettävässä muodossa (Hassija ym., 2024; Rodgers ym., 2023). Lisäksi päätöksentekoa automatisoivien järjestelmien tulisi tarjota selkeät perustelut valinnoilleen (Hassija ym., 2024). Bankins ym. (2022) korostavat, että oikeudenmukaisuuden kokemuksen kannalta on olennaista, että työntekijät saavat ymmärrettäviä syitä tehtyjen päätösten perusteista. Tämä voi tarkoittaa esimerkiksi perusteita esittelevien dokumenttien tarjoamista hylätyille työnhakijoille tai suoritusarvioiden selkeää avaamista työntekijöille.

HR-prosessien automatisoinnin rinnalla tulisi myös säilyttää mahdollisuus henkilökohtaiseen keskusteluun ja palautteeseen (Hassija ym., 2024). Mehrotra (2024) ja Hassija ym. (2024) painottavat, että työntekijöillä tulisi olla mahdollisuus esittää kysymyksiä ja saada lisätietoa tekoälypohjaisista päätöksistä, erityisesti päätösten vaikuttaessa heidän uraansa. Tämä voi tarkoittaa esimerkiksi sitä, että HR-asiantuntijat käyvät henkilökohtaisia keskusteluja työntekijöiden kanssa tekoälypohjaisista päätöksistä, ja että työntekijöillä on mahdollisuus valittaa epäoikeudenmukaisilta tuntuvista päätöksistä.

Empatian merkitys päätöksenteossa on myös keskeinen tekijä, joka tulee huomioida tekoälyn suunnittelussa ja käytössä (Bankins ym., 2022). Bankins ym. (2022) korostavat, että oikeudenmukaisuuden kokemukseen vaikuttaa se, kuinka inhimillisesti työntekijät kokevat tulevansa kohdelluiksi. HR-järjestelmien kehittäjien tulisi varmistaa, että järjestelmät tunnistavat työntekijöiden yksilölliset tilanteet. Tämä voi tarkoittaa esimerkiksi joustavampia suorituksen arviointikriteereitä tai mahdollisuutta henkilökohtaiseen näkökulmien huomiointiin tietyissä tilanteissa. Lisäksi työntekijöillä tulisi olla mahdollisuus haastaa tekoälyn valinnat ja saada niihin ihmisen tekemä uudelleenarviointi. Rodgers ym. (2023) ehdottavat, että HR-prosesseissa tulisi olla selkeä mekanismi, jonka kautta työntekijät voivat pyytää päätösten uudelleentarkastelua ja tarvittaessa saada lisäselvitystä niiden perusteista.

Organisaatioiden tulisi määritellä tekoälyn käyttöä koskevat eettiset periaatteet ja varmistaa, että HR-järjestelmät tukevat yrityksen arvoja ja monimuotoisuustavoitteita (Rodgers ym., 2023). Algoritmien tulee myös olla dynaamisia ja kehitettäviä, jotta ne voivat mukautua muuttuviin työympäristöihin ja lainsäädäntöön (Rodgers ym., 2023). Bankins ym. (2022) huomauttavat, että tekoälyjärjestelmien tulee sisältää mekanismeja, jotka mahdollistavat niiden päivityksen ja parantamisen, kun uusia haasteita tai epäkohtia havaitaan.

### 5.3 Yksityisyyden suojaaminen

Tekoälyn käyttö HR-prosesseissa on lisännyt tarvetta vahvistaa lainsäädäntöä ja sääntelyä, jotta työntekijöiden yksityisyys ja oikeudet turvataan (Bar-Gil ym., 2024). Euroopan Unionin AI Act on keskeinen lainsäädännöllinen kehys, joka asettaa tiukat vaatimukset tekoälyjärjestelmien turvallisuudelle ja eettisyydelle, mukaan lukien vaatimukset läpinäkyvyydestä, datan laadun varmistamisesta ja inhimillisestä valvonnasta (Hocken & King, 2023). AI Act:n sääntelyperiaatteet tukevat osittain Bar-Gil ym. (2024) näkemystä siitä, että nykyisiä tietosuojasäännöksiä, kuten EU:n yleistä tietosuojasetusta (GDPR), tulisi laajentaa työpaikoilla käytettäviin tekoälyjärjestelmiin. Tämä laajennus voisi tarkoittaa esimerkiksi sitä, että työnantajien on varmistettava tekoälyn suorittaman tietojenkerauksen ja -käytön perustuvan selkeään oikeudelliseen perustaan sekä työntekijän suostumukseen.

Vaikka AI Act pyrkii varmistamaan tekoälyn vastuullisen ja oikeudenmukaisen käytön muun muassa HR-prosesseissa, sääntelyn kehittämiseen liittyy vielä avoimia kysymyksiä. Esimerkiksi Roemmich ym. (2023) huomauttavat, että sääntely voisi asettaa tarkempia rajoituksia tekoälyn soveltamiselle, kuten emotion AI -teknologian käytölle työntekijöiden tunnetilojen analysoinnissa, jotta työntekijät eivät joutuisi jatkuvan tunnevalvonnan kohteiksi. Tämä näkökulma korostaa sääntelyn tarvetta tilanteissa, joissa tekoäly voi vaikuttaa työntekijöiden hyvinvointiin ja yksityisyyteen. Lisäksi Andrieux ym. (2024) ehdottavat, että työntekijöille tulisi taata oikeus kieltäytyä analysoinnista ilman kielteisiä vaikutuksia heidän urakehitykselleen. Näin ollen vaikka AI Act asettaa perustavanlaatuisia eettisiä ja oikeudellisia vaatimuksia, tutkijat painottavat, että sen soveltaminen HR-kontekstissa edellyttää lisäsääntelyä ja tarkennuksia työntekijöiden oikeuksien suojelemiseksi.

Bar-Gil ym. (2024) ehdottavat, että lainsäädäntöä voitaisiin lisäksi täydentää velvoittamalla yrityksiä ottamaan käyttöön tiukat tietoturva- ja yksityisyydensuojan minimistandardit tekoälyjärjestelmissään. Tämä voitaisiin toteuttaa esimerkiksi implementoimalla automaattisia anonymisointikäytäntöjä ja vaatimuksia tekoälypäästösten dokumentoinnille, jolloin työntekijät pystyvät tarkistamaan tietojensa käytön. Samalla työntekijöille tulisi antaa mahdollisuus hallita omaa dataansa esimerkiksi tarjoten selkeitä keinoja tarkastella, muokata tai poistaa heistä kerättyä tietoa.

Tekoälyn käytön riskienhallintaan ja vaikutustenarviointiin on kiinnitettävä erityistä huomiota, sillä sen vaikutukset eivät aina ole ennakoitavissa (Andrieux ym., 2024). Ro-

emmich ym. (2023) ehdottavat, että organisaatioiden tulisi ottaa käyttöön kattavat riskienhallintajärjestelmät sekä suorittaa säännöllisiä tietosuojaja- ja eettisiä vaikutustenarvioita (Privacy Impact Assessments, PIA ja Data Protection Impact Assessments, DPIA). Näiden järjestelmien avulla voidaan tunnistaa ja ehkäistä tekoälyn mahdollisia haittoja työntekijöiden yksityisyydelle ja autonomialle. Näihin arviointeihin sisältyy muun muassa tekoälypohjaisen päätöksenteon perusteiden, yksityisyyden riskien ja työntekijöiden vaikutusmahdollisuuksien tarkastelu.

Tekoälyn eettisen ja oikeudenmukaisen käytön edistäminen edellyttää myös HR-ammattilaisten koulutusta tietosuoja-asioissa (Rodgers ym., 2023). Rodgers ym. (2023) painottavat, että organisaatioiden tulisi järjestää säännöllisiä koulutuksia, joissa käydään läpi muun muassa tekoälyn soveltamista HR-prosesseissa, analysoitavan datan laatu, työntekijöiden tietosuojaan liittyvät oikeudet, sekä keinoja rajoittaa tietojen käyttöä tai valittaa epäoikeudenmukaisista päätöksistä. Erityisesti HR-ammattilaisten osaaminen on keskeistä, sillä heidän vastuullaan on varmistaa, että tekoälyn käyttö noudattaa organisaation eettisiä ja oikeudellisia periaatteita sekä edistää avoimuutta ja oikeudenmukaisuutta työpaikalla.

## **6 Johtopäätökset ja pohdinta**

Tutkielman päätavoitteena oli selvittää tekoälyn käytön eettisiä haasteita HR-prosesseissa ja esittää keinoja näiden riskien hallintaan. Työssä keskityttiin kolmeen keskeiseen osa-alueeseen: algoritmien puolueellisuuteen, päätöksenteon läpinäkyvyyden haasteisiin sekä työntekijöiden yksityisyyden suojaan. Näiden teemojen kautta työ pyrki tuomaan esiin tekoälyn käytön eettiset riskit ja niiden hallintaan liittyvät peruskysymykset. Tutkielmassa havaittiin, että eettisten haasteiden ratkaisut eivät aina ole yksiselitteisiä, ja useimmat kirjallisuudessa esitetyt keinot ovat pääasiassa teoreettisia ehdotuksia. Vaikka ratkaisuksi on kehitetty erilaisia periaatteita ja suosituksia, niiden käytännön toteutus voi usein osoittautua huomattavan moniulotteiseksi.

Tutkielma on onnistunut antamaan yleiskuvan olemassa olevasta tutkimuksesta, mutta sen rajallisen pituuden ja ei-empiristisen lähestymistavan vuoksi käytännön implementoinnin näkökulma joudutaan jättämään pintapuoliseksi. Tämän vuoksi tutkielmassa ei pystytä täysin tavoittamaan niitä syvällisiä haasteita ja konteksteja, joita organisaatioissa esiintyy tekoälyn käyttöönotossa. Myös jotkin aihealueet, joiden käsittely olisi

voinut olla vielä syvällisempää, on jätetty pois niiden valtavan laajuuden vuoksi, esimerkiksi kattavampi organisaatiopsykologian näkökulma ja juridiset sääntelykehys.

Tekoälyn sääntelyvaatimusten käsittely HR-prosesseissa osoittautui monitasoiseksi ilmiöksi. Nämä oikeudelliset vaatimukset muodostavat valtavan kokonaisuuden, joka ulottuu käytännön implementoinnista strategisiin päätöksiin ja sisäisiin kontrollimekanismeihin (Botero Arcila, 2024). Tällainen analyysi olisi vaatinut oman, laajempaan empiriseen ja teoreettiseen tutkimukseen perustuvan tutkimusarkkitehtuurin, joten sen käsittely jäi tämän työn puitteissa lyhyeksi.

Lähteiden osalta tutkimus perustuu ajankohtaiseen ja vertaisarvioituun kirjallisuuteen. Kuitenkin lähteiden kirjoa ja syvyyttä olisi voinut entisestään monipuolistaa. Olisi ollut hyödyllistä sisällyttää enemmän empiirisiä tutkimuksia ja tapaustutkimuksia, jotka käsittelevät tekoälyn käytön vaikutuksia käytännön HR-toimintaan. Lisäksi graafinen aineisto ja lainsäädännön analyysit olisivat voineet rikastaa kokonaiskuvaa ja tarjota konkreettisempaa dataa teorian tueksi. Kuitenkin lähteitä ja tutkimuksia aihealueen ympäriltä on tutkielman teon hetkellä vielä suppeasti aiheen ollessa uusi.

Edellä mainittujen puutteiden pohjalta herää useita jatkotutkimustarpeita. Tarvitaan empiirisiä tutkimuksia, jotka yhdistävät teoreettiset näkökulmat ja käytännön kokemukset, jotta voidaan arvioida ehdotettujen ratkaisujen toimivuutta reaaliaikaisissa HR-prosesseissa. Myös lähteissä nostetaan esiin joitain jatkotutkimustarpeita. Rodgers ym. (2023) esittävät, että erityisesti lainsäädännön, kuten EU:n AI Actin, vaikutusten ja käytännön soveltamisen analyysi voisi tarjota arvokasta tietoa siitä, miten juridinen kehys voisi tukea vastuullista tekoälyn hyödyntämistä työelämässä. Bankins ym. (2022) toteavat, että hybridimallit, joissa tekoäly tukee päätöksentekoa mutta lopullinen ratkaisu jää ihmiselle, voivat tasapainottaa algoritmien tehokkuuden ja inhimillisen harkinnan. He kuitenkin korostavat, että näiden mallien toimivuutta HR:n kontekstissa on vielä tutkittava tarkemmin. Jatkossa monitieteinen tutkimus, joka yhdistää tekoälytutkimuksen, organisaatiopsykologian ja oikeustieteen, on keskeisessä asemassa kehitettäessä eettisesti kestävää HR-päätöksentekoa.

On vielä valtavasti tarvetta jatkotutkimukselle, joka syventyy käytäntöön empiirisen tutkimuksen kautta. Nämä jatkotutkimukset ovat tärkeitä, jotta voidaan kehittää syvällisempiä ja käytännönläheisempiä ratkaisuja, jotka ottavat huomioon sekä teknologian nopean kehityksen että siihen liittyvän moninaisen sääntely-ympäristön.

## Lähdeluettelo

- Andrieux, P., Johnson, R. D., Sarabadani, J., & Van Slyke, C. (2024). Ethical considerations of generative AI-enabled human resource management. *Organizational Dynamics*, 53(1), <https://doi.org/10.1016/j.orgdyn.2024.101032>
- Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022). AI Decision Making with Dignity? Contrasting Workers' Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context. *Information Systems Frontiers*, 24(3), 857–875. <https://doi.org/10.1007/s10796-021-10223-8>
- Bar-Gil, O., Ron, T., & Czerniak, O. (2024). AI for the people? Embedding AI ethics in HR and people analytics projects. *Technology in Society*, 77, <https://doi.org/10.1016/j.techsoc.2024.102527>
- Bawa, S. S. (2023). How Business can use ERP and AI to become Intelligent Enterprise. *International Journal of Innovative Science and Research Technology*, 8, 8–11. <https://doi.org/10.5281/zenodo.7688737>
- Botero Arcila, B. (2024). AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight? *Computer Law & Security Review*, 54, <https://doi.org/10.1016/j.clsr.2024.106012>
- Devaraju, S. (2024). AI-Powered HRM and Finance Information Systems for Workforce Optimization and Employee Engagement. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 15(1), 269–281. <https://doi.org/10.61841/turcomat.v15i1.14940>
- Garg, S., Sinha, S., Kar, A. K., & Mani, M. (2021). A review of machine learning applications in human resource management. *International Journal of Productivity and Performance Management*, 71(5), 1590–1610. <https://doi.org/10.1108/IJPPM-08-2020-0427>
- Goswami, A. (2018). Human resource management and its importance for today's organizations. *Journal of Advances and Scholarly Researches in Allied Education*, 15(3), 128–135. <https://doi.org/10.29070/15/57308>
- Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, 16(1), 45–74. <https://doi.org/10.1007/s12559-023-10179-8>
- Hocken, E., & King, G. (2023). AI in the hiring process. *Strategic HR Review*, 22(3), 81–84. <https://doi.org/10.1108/SHR-03-2023-0014>

- Mehrotra, D. S. (2024). Challenges of Performance Management in the Era of Digital HRM. *International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR)*, <https://doi.org/10.31426/ijamsr.2024.7.10.7811>
- Nyathani, R. (2023). AI-Driven HR Analytics: Unleashing the Power of HR Data Management. *Journal of Technology and Systems*, *5*(2), <https://doi.org/10.47941/jts.1513>
- Okatta, C. G., Ajayi, F. A., & Olawale, O. (2024). Navigating the future: Integrating AI and machine learning in HR practices for a digital workforce. *Computer Science & IT Research Journal*, *5*(4), <https://doi.org/10.51594/csitrj.v5i4.1085>
- Ongsulee, P. (2017). Artificial intelligence, machine learning and deep learning. *2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE)*, 1–6. <https://doi.org/10.1109/ICTKE.2017.8259629>
- Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI & recruitment. *Computer Law & Security Review*, *53*, <https://doi.org/10.1016/j.clsr.2024.105966>
- Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review*, *33*(1), <https://doi.org/10.1016/j.hrmr.2022.100925>
- Roemmich, K., Schaub, F., & Andalibi, N. (2023). Emotion AI at Work: Implications for Workplace Surveillance, Emotional Labor, and Emotional Privacy. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–20. <https://doi.org/10.1145/3544548.3580950>
- Sehrawat, S. K. (2023). The Role of Artificial Intelligence in ERP Automation: State-of-the-Art and Future Directions. *Transactions on Latest Trends in Artificial Intelligence*, *4*(4). <https://www.ijsdcs.com/index.php/TLAI/article/view/474>
- Silva, M. S. A. e & Silva Lima, C. G. da S. (2017). The Role of Information Systems in Human Resource Management. *Teoksessa Management of Information Systems*. IntechOpen. <https://doi.org/10.5772/intechopen.79294>