

UAS-Journal 4/24 linkki: <https://urn.fi/URN:NBN:fi-fe2024112596653>

Tekoäly työyhteisöissä: yhdenvertaisuuden ja eettisyyden tarkastelua

Merja Drake

FT, yliopettaja, Haaga-Helia ammattikorkeakoulu

merja.drake@haaga-helia.fi

Eija Kärnä

KTT, lehtori, Haaga-Helia ammattikorkeakoulu

eija.karna@haaga-helia.fi

Anna Lahtinen

KTT, DBA, vanhempi tutkija, Haaga-Helia ammattikorkeakoulu

anna.lahtinen@haaga-helia.fi

Mariitta Rauhala

KTT, yliopettaja, Haaga-Helia ammattikorkeakoulu

mariitta.rauhala@haaga-helia.fi



Kuva 1. Työyhteisöt ovat tukevasti matkalla kohti tekoälyn tehokkaampaa käyttöä. Kuva Ksenia Senkova, Creative Commons Zero.

Tiivistelmä

Työyhteisöissä käydään vilkasta keskustelua tekoälyn hyödyntämisestä ja siihen liittyvästä osaamisesta. Organisaatiot ovat pilotoineet tekoälysovelluksia, sillä niistä on havaittu olevan hyötyä muun muassa prosessien tehostamisessa, laadun tasaamisessa ja tylsissä rutiinitehtävissä. Hyötyjen ohella tekoäly on aiheuttanut pelkoja työn menettämisestä, syrjinnästä ja saadun tiedon oikeellisuudesta. Tekoälyn syrjintä voi johtua koulutukseen käytetyn datan vinoumista ja yksipuolisuudesta ja tekoälykehittäjien miesvaltaisuudesta ja kehittäjien omista piilevistä vinoumista.

Artikkelissa esittelemme Tekoäly yhdenvertaisuuden edistäjänä -hankkeen tuloksia siitä, miten tekoälyä työyhteisöissä hyödynnetään, kuinka yhdenvertaisuuskysymykset näkyvät näissä teemoissa ja millainen rooli toimijuudella on tekoälymurroksessa.

Artikkelin aluksi kerromme tutkimuksen tavoitteista ja lyhyesti keskeistä käsitteistä kuten tekoäly, eettinen tekoäly, digitaalinen yhdenvertaisuus ja digitaalinen- ja kehittäjätoimijuus. Lopuksi avaamme alustavia tutkimustuloksia

Abstract

In work communities, there is lively discussion about the utilization of artificial intelligence and the skills related to it. Organizations have piloted AI applications, as they have been found to be beneficial in, among other things, streamlining processes, leveling quality, and handling tedious routine tasks. Alongside the benefits, AI has also caused fears of job loss, discrimination, and the accuracy of the information obtained. AI discrimination can result from biases and one-sidedness in the data used for training, the male dominance among AI developers, and the developers' own implicit biases.

In the article, we present the results of the "AI as a Promoter of Equity" project, focusing on how AI is utilized in work communities, how equity issues are reflected in these themes, and the role of agency in the AI revolution.

At the beginning of the article, we discuss the research objectives and briefly cover key concepts such as artificial intelligence, ethical AI, digital equality, and digital and developer agency. Finally, we reveal preliminary research results.

Ingressi

Tekoälystä keskustellaan runsaasti, mutta yhdenvertaisuuteen liittyviä näkökulmia on tuotu varsin vähän keskusteluun, vaikka tekoälyn syrjivyyttä on jossain määrin tunnistettu. Yhdenvertaisen työyhteisön näkökulmasta on tärkeää, että tekoälyn kehittämisestä, käyttöönotosta ja käyttötarkoituksista keskustellaan avoimesti työpaikoilla ja erilaisia näkemyksiä otetaan huomioon työprosesseja tekoälyllä kehitettäessä.

Halusimme siksi selvittää, kuinka yhdenvertaisuuden näkökulmat tulevat esille tietotyöntekijöiden tekoälyn pilotointi-, käyttö- ja kehittämiskontekstissa. Asiaa selvitimme kahdella tutkimuskysymyksellä: TK1 Kuinka tietotyön asiantuntijat hyödyntävät ja kehittävät tekoälyä, ja miten yhdenvertaisuuden kysymykset nousevat esiin näissä yhteyksissä? TK2: Kuinka tietotyön asiantuntijoiden ammatillinen toimijuus kehittyy digitoimijuudeksi ja kehittäjätoimijuudeksi tekoälyä käytettäessä ja kehittäessä?

Tutkimuksen aineisto hankittiin laadullisin ja määrällisin menetelmin ja toimintatutkimuksella. Helmikuusta maaliskuuhun 2024 auki olleeseen kyselyyn vastasi 469 vastaajaa. Kysely levitettiin erilaisten verkostojen, hankkeissa mukana olevien tietotyön ammattiliittojen ja työyhteisöjen välityksellä.

Kyselyä täydensivät laadulliset, temaattiset yksilö- ja ryhmähaastattelut ja kehittämistyöpajat. Haastatteluihin osallistui 44 henkilöä. Laadullinen aineisto analysoitiin sisällönanalyysillä. Sisällönanalyysillä

muodostetaan tutkittavasta ilmiöstä tiivistetty kuvaus (Tuomi & Sarajärvi 2009). Koodaus tehtiin Atlas.ti-ohjelmistolla Grounded Theoryn lähtökohdista (Charmaz 2006) Ylisassi ym. (2016) kehittämistoimijuuden tulkintakehikkoa hyödyntäen.

Tekoäly eettisen tekoälyn, yhdenvertaisuuden ja toimijuuden kontakteissa

Yhtäältä tekoälyä on kuvattu teknologiaksi, jonka avulla koneet voivat suorittaa monimutkaisia ja inhimillisiä taitoja (Sheikh ym. 2023). Toisaalta sitä on kuvattu laskentaa ja kognitiota hyödyntäväksi teknologiaksi, joka jäljittelee ihmisen oppimista, kanssakäymistä ympäristönsä kanssa, ratkoo ongelmia, tekee päätöksiä ja on luova. Valtava laskentateho auttaa suunnittelussa, kehittämisessä ja parhaisiin tuloksiin pääsemisessä. (Barr & Feigenbaum 2014; Boucher 2020; Feijoon & Kwon 2020.)

Yhdenvertaisuus merkitsee sitä, että tasa-arvoiseen lopputulokseen ei välttämättä päästä kohtelemalla kaikkia samalla tavalla. Yhdenvertaisten tulosten saavuttaminen edellyttävät resurssien ja mahdollisuuksien jakamista olosuhteiden ja tarpeen mukaan. (UN 2024.) Digitaalinen yhdenvertaisuus tarkoittaa yhtäläisiä mahdollisuuksia käyttää digitaalisia työkaluja, resursseja ja palveluja, lisätä digitaalista tietotaitoa, tietoisuutta ja taitoja (Davies ym. 2007; Gorski 2005).

Aiemmissa tutkimuksissa on havaittu huoli tekoälyn eettisyydestä, yksityisyyden ja organisaation tietosuojasta ja tekoälyn turvallisesta käytöstä, jotka hidastavat tekoälytuotteiden leviämistä (Banholzer ym. 2020; Borgesius 2018; Bostorm & Yudowsky 2014; Griffin ym. 2024). Tekoäly voi syrjiä nimen, sukupuolen tittelin tai etnisen tausta perusteella esimerkiksi rekrytoinnissa, lainojen myönnössä, verkkokaupoissa ja vankien vapauttamisessa (Courtland 2018; Crawford 2021; Du & Xie 2021; Dressel & Farid 2018; Hamilton 2018; O'Connor & Liu 2023; Vincent 2020; West ym. 2019). Eettinen tekoäly on läpinäkyvää, syrjimätöntä ja sen käyttöön liittyvät perusteet on hyvin määritelty (Borgesius 2020; EU 2019; Ojanen ym. 2022; Waelen & Wieczorek 2022; Young 2001).

Toimijuus on tärkeä elementti työn ja työyhteisön kehittämisessä. Toimijuutta voi esiintyä joko yksilön tai työyhteisön kollektiivisena ominaisuutena, johon sisältyvät ihmisen valinnat ja niiden seuraukset (Alasoini ym. 2022; Giddens 1984). Ammatilliseen toimijuuden ominaisuuksia ovat valta, osaaminen, ammatissa kehittyminen ja kokemus omaan työhön vaikuttamisesta ja kehittämisestä (Eteläpelto ym. 2014; Eteläpelto ym. 2017; Vähäsantanen ym. 2017).

Kehittämistoimijuuteen liittyy kyky kehittää omaa työtä, tuoda esiin uusia ideoita, edistää teknologioiden käyttöönottoa, käytettävyyttä ja motivaatio uusien sovellusten kehittämiseen (Emirbayer & Mische 1998; Haapasaari ym. 2016; Koivisto 2023; Ylisassi ym. 2016). Digitoimijuuteen puolestaan liittyy sekä digitaalisten mahdollisuuksien tunnistaminen että eettinen osaaminen ja teknologioiden vaikutusten hyötyjen ja haittojen ymmärtäminen. Digitoimijuuden rooli on merkittävä edistettäessä tasa-arvoa ja yhdenvertaisuutta, demokratiaa ja hyvinvointia. (Aagaard & Lund 2020; Alasoini ym. 2022; Koivisto 2023.)

Tietotyöntekijä käyttää tekoälyä monipuolisesti

Vastaajista 99 % tunnisti olevansa tietotyöntekijöitä. Heidän ikänsä vaihteli 30-vuotiaista 65-vuotiaisiin. Naisvastaajien määrä oli yli puolet. Vastaajien työyhteisöissä oli käytössä sekä omia tekoälysovelluksia että kaupallisia kuten CoPilot ja avoimia alustoja. Vastaajista noin 10 prosentilla ei ollut kokemusta tekoälyn käytöstä.

Tekoälyä käytettiin erityisesti tekstin ideointiin ja kielentarkastukseen, visualisointiin, suunnitteluun, raportointiin ja kielen kääntämiseen. Tekoälyosaamisessa liiketoiminnan ymmärrys, työprosessien kehittäminen, datan analysointi, käsittely ja visualisointi sekä tiimityöskentely katsottiin tärkeimmiksi osaamisiksi.

Yhdenvertaisuutta edistävät ja heikentävät tekijät

Tutkimusaineiston mukaan ihmiset tuntevat olevansa mukana työyhteisönsä tekoälykehityksessä hyvin eri tavoin. Vastauksissa oli sekä yhdenvertaisuutta edistäviä että heikentäviä tekijöitä. Edistävinä tekijöinä mainittiin muun muassa, että työyhteisössä otetaan kaikki tasavertaisesti mukaan pilotointiin, tietoa jaetaan omissa kanavissa, viestintä on avointa ja tasa-arvoista kielitaidosta huolimatta, jokaisella on mahdollisuus tehdä erilaisia omia kokeiluja ja että työyhteisössä on omat tekoälyä koskevat eettiset periaatteet.

”Edistävintä on se, että kaikki on otettu mukaan testailuun ja jakamaan kokemuksia tekoälystä.”

”Omassa organisaatiossa toteutuu yhdenvertaisuus avoimen viestinnän kautta ja on mahdollista osallistua kehittämiseen.”

Yhdenvertaisuutta heikentävinä tekijöinä mainittiin huoli iäkkäämpien henkilöiden mahdollisuudesta testata ja pilotoida. Tekoälyn pilotointi ja kehittäminen katsottiin olevan pienten piirien asia, mistä ei saatu riittävästi tietoa ja että työyhteisöllä ei ollut selkeitä tavoitteita tekoälyn käytölle ja hyödyntämiselle. Kehittämisen ja käyttöönoton pitäisi siten olla huomattavasti kokonaisvaltaisempaa ja läpinäkyvämpää.

”Vanhempia työntekijöitä ei ainakaan oteta huomioon, eikä heidän työnsä kehittämisestä olla kiinnostuneita.”

”Heikentävänä ilmiönä ehkä iäkkäämpien työntekijöiden pelko siitä, että robotti vie heiltä työt.”

”Tekoälyä kehitetään ehkä tietohallinnossa ja joissain tiettyjen yksiköiden prosesseissa. Näistä ei ole talotasolla juurikaan tietoa.”

Yhdenvertaisuuden toteutumisesta oltiin huolestuneita, koska kaikilla ei nähty olevan samanlaisia valmiuksia opiskella ja omaksua asioita. Koulutuksen saatavuutta pidettiin myös riittämättömänä. Vastaajat kokivat, että osa työntekijöistä jää kokonaan tekoälykeskustelun ja käytön ulkopuolelle. Uhkakuvana nähtiin töiden siirtyminen tekoälylle, mikä johtaa irtisanomisiin. Epätietoisuus siitä, kenen työpaikat ovat vaarassa tuntui aiheuttavan eettistä stressiä.

Vastaajissa oli myös niitä, joiden mukaan yhdenvertaisuutta ei tule sekoittaa tekoälyn kehittämiseen tai hyödyntämiseen lainkaan. Yllätys oli, että yhdenvertaisuuden käsite oli työyhteisössä melko tuntematon.

”Tekoälyn hyödyntämisessä ei tarvitse ottaa eikä pidä ottaa huomioon mitään yhdenvertaisuusnäkökulmia.”

”Koko työyhteisö ei tunne yhdenvertaisuuden käsitettä”

Vastaajat olivat innokkaita, kriittisiäkin tekoälykokeilijoita, jotka jakoivat mielellään kokemuksiaan omille kollegoilleen. Selkeitä pilotointien jälkeisiä toimintasuunnitelmia kuitenkin kaivattiin. Osa olisi halunnut olla myös selkeämmin mukana kehittämässä tekoälyllä omia työprosessejaan. Heillä oli halu olla tekoälymuutoksen kehittäjä- ja digitoimijoita.



Kuva 2: Latu kohti kehittäjä- ja digitoimijuutta on avattu, mutta lisää osallistujia tarvitaan, jotta tekoälyn käyttö ja kehittämien saadaan yhdenvertaiseksi. Kuva Roine Piirainen, Creative Commons Zero.

Kohti kehittäjä- ja digitoimijuutta

Tulokset osoittavat, että tekoäly vaikuttaa merkittävästi tietotyön asiantuntijoiden toimintaan ja että sen käyttöönotto ja kehittäminen nostavat esiin monipuolisesti yhdenvertaisuuden kysymyksiä. Työyhteisöissä on tarvetta uudentylaiselle toimijuudelle, joka huomioi sekä teknologian mahdollisuudet että sen tuomat eettiset haasteet. Kehittäjätoimijuus ja digitoimijuus ovat avainasemassa näiden muutosten edistämisessä.

Kehittäjätoimijuuden avulla asiantuntijat voivat aktiivisesti osallistua tekoälyn suunnitteluun ja käyttöönottoon. Tämä vaatii kykyä tunnistaa ja ratkaista teknologian käyttöön liittyviä eettisiä ja yhdenvertaisuushaasteita sekä tuoda esiin aloitteita ja uusia ideoita työyhteisön arjessa. Digitoimijuus puolestaan korostaa tekoälyratkaisujen ymmärtämistä ja kehittämistä eettisesti kestäväällä tavalla. Se tarkoittaa myös valmiuksia arvioida ja soveltaa tekoälyä koko työyhteisöä palvelevaksi.

Työyhteisöjen tulisi aktiivisesti edistää kehittäjä- ja digitoimijuuden kehittymistä tarjoamalla koulutusta, joka tukee teknologian monipuolista ja eettistä käyttöä. Näin työntekijät voivat kehittää omaa työtään ja kokea vaikuttavansa siihen, miten teknologiaa hyödynnetään työpaikoilla. Organisaatioiden tulee luoda tilaa avoimelle keskustelulle tekoälyn käytön riskeistä ja mahdollisuuksista, jotta kaikki työntekijät tuntevat olevansa osa kehitystä.

Tekoälyn avulla voidaan luoda yhdenvertaisempia ja oikeudenmukaisempia työyhteisöjä, kun sen käyttöä ja kehittämistä ohjataan eettisesti ja vastuullisesti. Näin varmistetaan, että teknologinen kehitys palvelee kaikkia työntekijöitä ja organisaation pitkän aikavälin tavoitteita.

Huomautukset:

Artikkeli perustuu Haaga-Helian ammattikorkeakoulun toteuttaman ja Työsuojelurahaston osittain tukeman "AIE – Tekoäly yhdenvertaisuuden edistäjänä" -hankkeen alustaviin tutkimustuloksiin.

Asiasanat:

Tekoäly, eettinen tekoäly, yhdenvertaisuus, digitaalinen yhdenvertaisuus, digitaalinen toimijuus, kehittäjätoimijuus

Lähteet:

Aagaard, T. & Lund, A. 2020. Digital Agency in Higher Education: Transforming Teaching and Learning. London and New York: Routledge.

Alasoini, T., Ala-Laurinaho, A., Käsälä, M., Saari, E., & Seppänen, L. 2022. Työelämän digikuilujen yli: digitalisaatio kaikkien kaveriksi. Työterveyslaitos.

<https://www.julkari.fi/bitstream/handle/10024/143939/TTL-978-952-261-997-6.pdf?sequence=1&isAllowed=y>.

Banholzer, M., Fletcher, B., LaBerge, L. & McClain 2023. Companies with innovative cultures have a big edge with generative AI. <https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/companies-with-> Companies with innovative cultures have a big edge with generative AI.

Barr, A. & Feigenbaum, E.A. 2014. Artificial intelligence. In The handbook of artificial intelligence, vol 1. (eds.) Barr & Feigenbaum. Butterworth-Heinemann.

Borgesius, F.Z. (2018). Discrimination, artificial intelligence, and algorithmic decision making. Strasbourg: Council of Europe.

Borgesius, F.Z. (2020.) Strengthening legal protection against discrimination by algorithms and artificial intelligence. *The International Journal of Human Rights*. Vol 24: 10. 1572-1593.

Boström, N. & Yuodowsky, E. 2014. The ethics of Artificial Intelligence, in *The Cambridge Handbook of Artificial Intelligence*, pp. 316 – 334 DOI: <https://doi.org/10.1017/CBO9781139046855.020>.

Boucher, P. 2020. Artificial intelligence: How does it work, why does it matter, and what can we do about it? EPRS | European Parliamentary Research Service.

Charmaz, C. 2006. *Constructing Grounded Theory – A Practical Guide Through Qualitative Research*. Sage Publications. London. UK.

Courtland, R. 2018. Bias detectives: The Research striving to make algorithms fair. *Nature* **558**, 357-360. doi: <https://doi.org/10.1038/d41586-018-05469-3>

Crawford, K. 2021. *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Davies, T. Fueller, M. Jackson, S., Pitman, J. & Sweet, J. 2007. A National Consideration of Digital Equity. Washington C.D. International Society for Technology and Education. <<http://www.iste.org.digitalequity>>

Dressel, J. & Farid, H. 2018. The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4 (1) (2018), p. eaao5580, [10.1126/sciadv.aao5580](https://doi.org/10.1126/sciadv.aao5580)

Du, S. & Xie C. 2021. Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research* 129 pp. 961–974.

Emirbayer, M., & Mische, A. 1998. What is agency? *American Journal of Sociology*, 103(4), 962–1023. <https://doi.org/10.1086/231294>.

Eteläpelto, A., Vähäsantanen, K., Hökkä, P. & Paloniemi, S. 2014. Miten käsitteellistää ammatillista toimijuutta työssä? *Aikuiskasvatus* 34 (3), 202–214. <https://doi.org/10.33336/aik.94100>.

Eteläpelto, A., Vähäsantanen, K., Hökkä, P. & ja Paloniemi, S. 2017. Tutkimus- ja kehittämishankeen tausta ja lähtökohdat. Teoksessa K. Vähäsantanen, S. Paloniemi, P. Hökkä, P. & A. Eteläpelto (toim.) *Ammatillinen toimijuus: rakenne, mittari ja tuki*. Jyväskylä: Jyväskylän yliopisto, 5–13. <http://urn.fi/URN:ISBN:978-951-39-6980-6>.

European commission AI HLEG 2019. Ethics guidelines for trustworthy AI. Brussels.

Feijoo, C. & Kwon, Y. 2020. AI impacts on economy and society: Latest developments, open issues and new policy measures. *Telecommunication Policy* 44.

Giddens, A. 1984. *The constitution of society: Outline of the theory of structuration*. Polity Press.

Griffin, T., Green, P. & Weile, J. 2024. The ethical agency of AI developers. *AI & Ethics*. Volume 4, pp. 179–188.

Gorski, P. 2005. Education equity and the digital divide. *AACE Journal*, 13(1), 3-45. Retrieved from <https://psycnet.apa.org/doi/10.1037/0022-3514.44.1.113> Gorski, P.C. (2009). Insisting on digital equity: Reframing the dominant discourse on multicultural education and technology. *Urban Education*, 44(3), 348–364.

- Hamilton, I. 2018. Why it's totally unsurprising that Amazon's recruitment AI was biased against women. <
<https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10>>.
- Horvitz, E & Mitchell, T. 2024. Scientific Progress in Artificial Intelligence: History, Status, and Futures. In: *Realizing the Promise and Minimizing the Perils of Artificial Intelligence for the Scientific Community*. Eds Kathleen Hall Jamieson, Anne-Marie Mazza, and William Kearney. University of Pennsylvania Press.
- Koivisto, T. 2023. Digitoimijuus Terveysthuollon ammattilaisen työssä. Tampereen yliopisto.
- O'Connor, S. & Liu, H. 2023. Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities. *AI & Society*, Springer.
- Ojanen, A., Sahlgren, O., Vaiste, J., Björk, A. Mikkonen, J., Kimppa, K., Laitinen, A. & Oljakka, N. 2022. Algoritminen syrjintä ja yhdenvertaisuuden edistäminen: Arviointikehikko syrjimättömälle tekoälylle. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2022:54.
<https://julkaisut.valtioneuvosto.fi/handle/10024/164290>.
- Seppänen, A. & Heliste, A. 2022. Eettisen digitalisaation kyvykkyyksiä rakentamassa. Organisaation eettisen toimintakyvyn kehittäminen Goforella.
- Sheikh, H., Prins, C. & Schrijvers, E. 2023. Mission AI. The New System Technology. Springer.
- Tuomi, J. & Sarajärvi, A. 2009. Laadullinen tutkimus ja sisällönanalyysi. Jyväskylä: Tammi.
- UN 2024. Sustainable Development Report 2024. Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (E/CN.3/2017/2).
- Vincent, J. 2020. Services that uses AI to identify gender based on names looks incredibly biased.
<https://www.theverge.com/2020/7/29/21346310/ai-service-gender-verification-identification-genderify>.
- Vähäsantanen, K., Paloniemi, S., Räikkönen, E., Hökkä, P. & Eteläpelto, A. 2017. Ammatillisen toimijuuden moniulotteinen rakenne ja mittarikehittely. Teoksessa K. Vähäsantanen, S. Paloniemi, P. Hökkä, P. & A. Eteläpelto (toim.) Ammatillinen toimijuus: rakenne, mittari ja tuki. Jyväskylä: Jyväskylän yliopisto, 13–33.
<http://urn.fi/URN:ISBN:978-951-39-6980-6>.
- Waelen, R. & Wiecek, M. 2022. The struggle for AI's Recognition: Understanding the Normative Implications of Gender Bias in AI with Honneth's Theory Recognition.
- West, S.M., Whittaker, M. & Crawford, K. 2019. Discriminating Systems: Gender, Race and Power in AI. AI Now Institute. Retrieved from <https://ainowinstitute.org/discriminatingystems.html>.
<https://ainowinstitute.org/wp-content/uploads/2023/04/discriminatingystems.pdf>
- Ylisassi, H., Hasu, M., Heikkilä, H., Käpykangas, S., Saari, E., Seppänen, L. & Valtanen, E. 2016. Työntekijöiden kehittämistoimijuutta edistämässä. Kehittämismenetelmäkokeilujen tuloksia vanhuspalveluissa.
- Young, I. M. (2001). Equality of whom? Social groups and judgments of injustice. *Journal of Political Philosophy*, 9(1), 1–18. <https://doi.org/10.1111/1467-9760.00115>.