Mida tähendab „õiglane“ andmestunud heaoluriigis?



MARIS MÄNNISTE
Tartu Ülikooli kriitiliste andmeuuringute lektor



LEE RAUDSEPP
Tartu Ülikooli infoõiguse nooremteadur

Kujutage ette inimest, kes astub pärast koondamist Töötukassa uksest sisse. Ta vastab küsimustele oma hariduse, töökogemuse, keeleoskuse ja juhiloa kohta. Nõustajal on lisaks vestluses saadud teabele kasutada andmepõhine prognoos selle kohta, kui tõenäoliselt inimene töö leiab. See hinnang võib mõjutada seda, kuidas nõustaja inimese abivajadust tõlgendab ja talle tuge kavandab. Prognoos võib aga jääda inimesele endale nähtamatuks. Nii võib tema kohta tehtavat otsust mõjutada hinnang, mille kujunemist ja rolli ta ei tunne.

Üks heaoluriigi keskseid küsimusi on, kelle vajadusi ja millisel määral ühiste vahendite arvelt toetada. Teenuste kavandamisel toetavad nõustajate otsuseid üha sagedamini algoritmilised hinnangud, mis prognoosivad näiteks inimese töö leidmise tõenäosust. Sellisest prognoosist üksi ei selgu, kui palju ja millist abi inimene vajab. Viimasel kümnendil on paljud Euroopa avalike tööturuteenuste osutajad võtnud kasutusele erinevaid andmepõhiseid otsustustoe tööriistu, et piiratud nõustamisressurssi klientide vahel jaotada (Allhutter jt, 2020; Körtner & Bonoli, 2023; Delpierre jt, 2024; Pultz, 2025; Berman jt, 2024). See on tekitanud elava diskussiooni selle üle, kas seesugune lähenemine on õiglane.

Siinkohal olekski paslik küsida, mida tähendab algoritmide kontekstis „õiglane“? Kelle jaoks peab algoritm õiglane olema ning kelle käes on võim neid vastuseid kujundada? Esitatud küsimused puudutavad laiemalt seda, kuidas mõistame õiglast abi andmestunud heaoluriigis. Avalike teenuste korraldamisel tuleb otsustada, kuidas jagada piiratud ressursse, milliseid vajadusi pidada esmatähtsaks ning kui palju arvestada inimeste erinevate olukordadega. Andmepõhiste otsustustugede kasutamine annab põhjust uurida, kuidas need valikud väljenduvad andmetes, hindamiskriteeriumides ja igapäevases nõustamistöös. Milline arusaam õiglasest abist neis kinnistub? Kelle teadmisi inimese olukorra kohta arvestatakse? Ning milline võimalus on teenuse kasutajal enda kohta kujunevat hinnangut mõista ja mõjutada?

Nende küsimuste arutamiseks toetume rahvusvahelises AI FORA projektis korraldatud mängulisele katsetusele, mis lähtus Eesti Töötukassa otsustustoe OTT kasutamist käsitlenud uuringutest (Männiste jt, 2026). See kogemus pakub lähtekoha, et arutleda, kuidas muuta hindamise alused ja tagajärjed inimestele nähtavaks ning luua võimalusi nende üle kaasa rääkida. Tähelepanu all on ka see, mida seesugune osalemine võimaldab ja millised küsimused jäävad selle abil lahendamata.

Andmestunud heaoluriik ja õigluse dilemma

Heaoluriigid on pikka aega süstemaatiliselt ja suures mahus kogunud oma elanike kohta mitmesuguseid andmeid ning sellest on saanud riigivalitsemise lahutamatu osa (Dencik & Kaun, 2020). Digitaliseerimine ja andmestumine on seda protsessi märkimisväärselt kiirendanud. Järjest enam valdkondi, elutegevusi ja isikuomadusi numbristatakse, kvantifitseeritakse ning jälgitakse (Kaun & Masso, 2025). Kaun ja Masso väidavad, et andmepõhiste tehnoloogiate kasutuselevõtt avalikus sektoris ongi viinud meid andmestunud heaoluriikide (ingl *data welfare state*) tekkeni ehk riikideni, kus andmed on ühtaegu abivahend, kütus ja otsustaja ning nende roll sõltub poliitilisest ja kultuurilisest kontekstist.

Riigid kasutavad avalikus sektoris üha enam automatiseeritud otsustusmehhanisme (ingl *automated decision-making* (ADM)) (vt nt Chiusi jt, 2020; Misuraca & van Noordt, 2020). Peamiselt võetakse ADM-e kasutusele kahel põhjusel: esiteks soovitakse suurendada tõhusust, parandada sidusust ja võimaldada käsitleda rohkem juhtumeid (Veale & Brass, 2019) ning teisalt soovitakse organisatsioonis kulusid vähendada ja administratiivseid protsesse lihtsustada (Valli Buttow & Weerts, 2022). Tuleb aga arvestada, et algoritmidele otsustusõiguse delegeerimine ei ole kõrvalmõjudeta. Uuringutes ja kodanikuühiskonna organisatsioonide raportites on juhitud tähelepanu algoritmilisele kallutatusele, ebaselgele vastutusele, vähesele läbipaistvusele ning otsuste kvaliteedile, eriti sotsiaalvaldkonnas (nt Alston, 2019; Eubanks, 2018; Allhutter jt, 2020; Kaun, 2022; Rinta-Kahila jt, 2024; Amnesty International, 2024).

Lisaks muudab selliste süsteemide kasutuselevõtt andmestunud heaoluriigis inimese ja riigi vahelist võimusuhet. Jørgensen (2021) osutab, et intensiivne andmestumine võib süvendada inimese ja riigi vahelist info- ja võimu asümmeetriat ning seda ohtu võivad võimendada optimistlikud käsitlused andmepõhise valitsemise võimalustest. Koos sellega kerkib küsimus, kellel on õigus määratleda, millised andmed ja hindamiskriteeriumid peaks olema inimese abivajaduse hindamise aluseks. Kui need valikud jäävad üksnes süsteemi arendava ja kasutava organisatsiooni teha, võib inimese enda arusaam oma olukorrast jääda otsustamisel tagaplaanile. Seetõttu on oluline küsida, millised võimalused on inimesel mõista ja mõjutada seda, kuidas tema andmeid teda puudutavate otsuste tegemisel kasutatakse.

Üks keskseid küsimusi algoritmiliste süsteemide riskide maandamisel on õiglus (Starke jt, 2022), mis on usaldusväärse tehisintellekti põhimõtte alus (OECD, 2019). Algoritmilise õigluse käsitlused pööravad tähelepanu sellele, et automatiseeritud otsused ei põhjustaks diskrimineerimist ega muid ebaõiglasi tagajärgi (Mitchell jt, 2021). Erinev kohtlemine ei tähenda tingimata ebaõiglust, sest suurem abivajadus võib õigustada ka suuremat toetust. Arusaamad õiglusest ja selle rakendamisest sõltuvad poliitilisest raamistikust, kultuurilistest hoiakutest ning valdkonnast, kus süsteemi kasutatakse (Starke jt, 2022; Kaun jt, 2024). Nii nagu tõdevad Ahrweiler jt (2024), ei olemas olemas lähenemist, mida tajutaks kõikjal õiglasena. Õigluse tajumist tuleb seejuures eristada otsuste tegelike tagajärgede hindamisest: õiglasena tunduv süsteem võib siiski mõnd rühma ebasoodsalt kohelda (Starke jt, 2022).

Õiglusega seotud pingeid ilmestavad mitmed Euroopa heaoluriikides kasutusele

võetud süsteemid. Näiteks Hollandi toetuspettuste tuvastamise süsteem SyRI ühendas avalike asutuste andmeid, et koostada inimeste kohta riskiprofiile. Süsteemi kasutati peamiselt piirkondades, kus elas palju väikese sissetulekuga inimesi, sisserändajaid ja etnilisi vähemusi, mistõttu tõstatati küsimus selle rollist ebavõrdse kohtlemise süvendamisel. Inimesed ei olnud alati teadlikud, et neid hinnatakse, ning riskiprofiilide koostamise alused jäid läbipaistmatuks. SyRI tunnistati 2020. aastal kohtus õigusvastaseks (Appelman jt, 2021). See juhtum osutab, kui oluline on uurida, kellele süsteemi rakendamine keskendub ning millised võimalused on inimestel enda kohta kujundatud hinnangut mõista ja vaidlustada.

Õigluse küsimus puudutab ka seda, milliste eesmärkide saavutamisse heaoluriik oma ressursse suunab. Näiteks Soome sotsiaalkindlustusamet Kela on arendanud tehnoloogilisi võimalusi, et toetuspettuseid tuvastada, kuigi kahtlustatud väärkasutuse rahaline maht on väljamakstud toetustega võrreldes väike (vt van Gerven & van Lancker, 2026). See võrdlus ei näita iseenesest kontrollimeetmete tulemuslikkust, kuid annab põhjust arutada nende proportsionaalsuse üle. See tõstatab küsimuse, kuidas tasakaalustada avalike vahendite kasutamise kontrollimist ning inimeste toetamist ja nende probleemide algpõhjustega tegelemist.

Need näited puudutavad eelkõige pettuste tuvastamist. Töötute nõustamise kontekstiga ühendab neid küsimus, kuidas andmepõhised hinnangud mõjutavad inimese kohtlemist ja avalike ressursside jaotamist. Nõustamisel tuleb näiteks otsustada, kas on põhjendatud pakkuda suuremat tuge, kui töö leidmise tõenäosus on prognoosi järgi väike, või suunatakse piiratud ressursid inimestele, kelle puhul oodatakse kiiremat tulemust. Õigluse hindamiseks tuleks seega vaadelda nii hinnangu kujunemist kui ka põhimõtteid, mille järgi selle alusel tegutsetakse.

Õigluse tajumist käsitlev kirjandus toob esile võimaluse otsuseid vaidlustada ja saada nende kohta selgitusi, kuigi selgituste mõju sõltub olukorrast (Starke jt, 2022). Sun ja Tang (2021) osutavad tarbijate kontekstis ka tajutud kontrolli olulisusele. Töötute nõustamisel suunab see küsima, millised võimalused on inimesel enda kohta tehtavat hinnangut mõjutada. Siinses arutelus lisandub küsimus, kuivõrd võimaldab otsustusprotsess arvestada inimese olukorra nende tahkudega, mida olemasolevad andmed ja mõõdetavad tunnused ei kajasta. Siin on tähtis nii inimese võimalus oma olukorda selgitada kui ka nõustaja võimalus seda teavet otsustamisel arvesse võtta. Seega sõltub õiglus ka sellest, kuidas arvestatakse algoritmi hinnangut, nõustaja erialateadmisi ja inimese enda kogemust.

Erinev kohtlemine ei tähenda tingimata ebaõiglust, sest suurem abivajadus võib õigustada ka suuremat toetust.

Mäng kui uurimismeetod

Algoritmilise hindamise õigluse uurimise ühe võimalusena käsitleme AI FORA projektis korraldatud mängulist katsetust, mida on kirjeldanud Männiste, Vihalemm ja Trumm (2026). Selle lähtekoht oli Eesti Töötukassa otsustustugi OTT, mis prognoosib tööotsija töö leidmise tõenäosust ja toetab nõustaja otsuseid. Varasemad uuringud näitasid, et nõustajad tõlgendavad OTT-i hinnanguid koos oma professionaalsete teadmiste ja kliendi olukorda puudutava teabega (Vihalemm jt, 2025; Weitz jt, 2024).

Teenuse kasutajale võivad otsustustoe roll ja hindamise alused jääda nähtamatuks. Selleks et muuta need arutletavaks

ka inimestele, kes tööriista ise ei kasuta, rakendati uuringus tõsiste mängude (ingl *serious games*) lähenemist. See võimaldas osalejatel võtta nõustaja või tööotsija roll, katsetada erinevaid hindamisreegleid ning arutleda nende tagajärgede üle. Mängus kasutati OTT-ist inspireeritud lihtsustatud hindamisreegleid; tegelikku OTT-i mudelit selles ei rakendatud.

2024. aasta oktoobris toimunud katsetuses osales 18 Tartu Ülikooli sotsiaalteaduste magistranti. Nad olid varem tutvunud tehisintellekti ühiskondlike mõjudega, kuid neil puudus kogemus Töötukassa teenustega. Osalejatest 15 mängisid tööotsijaid ja kolm nõustajaid. Tööotsijate rollid lähtusid kolmest profiilist: noor tööotsija, raskete terviseprobleemidega tööotsija ja mitme ülalpeetavaga tööotsija. Uurijad arvutasid mängu hindamisreeglite alusel skoorid ning edastasid need nõustajatele.

Mäng toimus kolmes ringis. Esimeses ringis ei teadnud osalejad skoori arvutamise reegleid; neid selgitati sellele järgnenud arutelus. Järgmisteks ringideks said osalejad valida muudatusi etteantud loetelust ning pidid neis ühiselt kokku leppima. Uuringus analüüsiti osalejate kirjalikke päevikuid ja arutelusid, keskendudes sellele, milliseid hindamise ja toe jaotamise põhimõtteid peeti õiglaseks, milliseid asjaolusid sooviti arvesse võtta ning kuidas reeglimuudatusi põhjendati.

Algoritm kui poliitiline otsus õiglusest

OTT-ist inspireeritud mäng tõi esile, kuidas õigluse tähendus kujuneb konkreetsete hindamis- ja jaotamisotsuste kaudu. Tervise ja hoolduskohustuste lisamine hindamisse võimaldas tööotsijate olukorda üksikasjalikumalt kirjeldada. Samal ajal hakkasid kliendirollis osalejad rohkem rõhutama oma tegelase vajadusi ning tähelepanu ühise nõustamisressursi jaotamisele vähenes. See tõi nähtavale pinge kahe põhjendatud ootuse vahel: inimene soovib, et tema eripäraga arvestataks, kuid asutus peab jagama piiratud tuge kõigi abivajajate vahel. Rohkem teavet üksikisiku kohta ei lahenda iseenesest küsimust, kuidas neid ootusi tasakaalustada. Õiglus ei sõltu üksnes algoritmi parameetritest, vaid ka sellest, milliseid vajadusi peetakse oluliseks ja millistel põhimõtetel toe jagamises kokku lepitakse.

See seob algoritmilise hindamise heaoluriigi solidaarsuse ja ühiste ressursside kasutamisega. Näiteks võib väike prognoositav töö leidmise tõenäosus anda põhjuse inimesele rohkem tuge pakkuda. Samas võidakse piiratud ressursid suunata neile, kelle puhul oodatakse kiiremat tulemust. Mõlemal juhul võib lähtehinnang olla sama, kuid rakendatav abi jaotamise põhimõte erineb. Nagu pettuste tuvastamise näidete puhul, sõltub hinnangu tähendus siin sellest, millist eesmärki asutus taotleb ja kuidas ta selle saavutamiseks ressursse jaotab. Üksikasjalikule individuaalsele hindamisele keskendumine võib jätta tagaplaanile ka struktuurse ebavõrdsuse (Allhutter jt, 2020; Mann, 2020). Mängus tõstatatud küsimus asutuste pädevuse piiridest avab selle praktilise külje: terviseprobleemi või hoolduskohustuse arvestamine hindamisel ei taga veel inimesele sobivat abi. Selleks on vaja teenuseid, ressursse ja eri valdkondade koostööd. Digitaalsete lahenduste kasutuselevõtt annab seega uue kuju heaoluriigi senistele poliitilistele küsimustele: keda toetada, millistel alustel ja milliste vahenditega?

Õiglust kujundab ka see, kuidas hindamisreegleid igapäevases töös rakendatakse. Mängus võisid sarnase profiiliga tööotsijad saada erineva tulemuse, sest nõustaja rollis osalejad hindasid näiteks terviseprobleemi raskust erinevalt. Arvulise skoori taga olid seega inimlikud tõlgendused. Samal ajal soovisid osalejad, et nõustajatel oleks rohkem võimalusi lisaküsimusi esitada ja hindamist inimese olukorraga kohandada. Siin ilmneb ühetaolise kohtlemise ja kaalutlusruumi vaheline pinge: erandite tegemine võib aidata inimese vajadusi paremini arvestada, kuid võib tuua kaasa

ka põhjendamatuid erinevusi. See seostub otsustusprotsessi õiglusega ning varasemas kirjanduses käsitletud võimalustega otsuseid põhjendada ja vaidlustada (Starke jt, 2022). Seetõttu vajab inimlik otsustamine ühiseid põhimõtteid, otsuste põhjendamist ja nende läbivaatamise võimalust. Vastutus tuleb määratleda süsteemi arendavate ja kasutavate inimeste ning organisatsioonide vahel, sealhulgas selgitada, millal saab nõustaja algoritmi hinnangust kõrvale kalduda ja kuidas ta seda põhjendab.

Osalejate aruteludes kerkis esile ka see, et hinnang sõltub inimese valmisolekust oma olukorda avada. See suunab tähelepanu küsimusele, kas sobiva toe saamine võib hakata sõltuma inimese oskusest oma vajadusi hindamissüsteemile arusaadavalt esitada. Osalejad pidasid hindamise läbipaistvust oluliseks, kuid väljendasid ka kartust, et reeglite tundmine võimaldab süsteemi enda kasuks mõjutada. See oli osalejate tõstatatud võimalus, mille kõrval tuleb eristada tahtlikku eksitamist inimese põhjendatud soovist oma olukorda selgitada ja enda eest seista. Tehniliselt avatud süsteem võib jääda inimese jaoks raskesti mõistetavaks, kui talle pole seda arusaadavalt selgitatud ega antud võimalust hinnangut mõjutada. Nii seostub „musta kasti“ küsimus ka sellega, kellel on tegelikult võimalus otsustusprotsessis osaleda (Pasquale, 2015). Jørgenseni (2021) kirjeldatud teabe- ja võimuasümmeetria seisukohalt on seega oluline, kas inimene saab lisaks reeglitega tutvumisele oma andmeid täpsustada ja hinnangut vaidlustada.

Siin avaldub mängulise uurimise väärtus. Ühine otsustusolukord võimaldas osalejatel hindamisreegleid läbi proovida, nende tagajärgi arutada ja muudatusi põhjendada. Nii muutusid nähtavaks erinevad arusaamad õiglasest abist ning nendevahelised pinged. Mängus osalenud magistrantide arutelud annavad lähtekoha, et neid küsimusi koos tegelike teenusekasutajate ja nõustajatega edasi uurida. Andmestunud heaoluriigi kujundamisel vajabki tähelepanu see, kelle teadmised ja väärtushinnangud jõuavad hindamisreeglitesse, kes vastutab nende rakendamise eest ning kuidas saavad reegleid mõjutada inimesed, kelle elu nende alusel tehtud otsused puudutavad.

Kokkuvõttes ei peaks me küsima üksnes seda, kas algoritm on õiglane, vaid ka seda, kelle arusaamast õiglusest me lähtume ja millistes otsustusprotsessides on inimese hinnang olulisem kui tabeliandmed. Sellistele teravatele, kuid keerukatele küsimustele vastamine on järgmise kümnendi üks tähtsamaid ülesandeid nii heaoluriikide parlamentidele kui ka ühiskondadele. Meeles tuleb pidada sedagi, et tegu ei ole retooriliste küsimustega.

KASUTATUD ALLIKAD

AHRWEILER, P., GILBERT, N., JURANYI, Z., BICET, M., COLL, A. S., KAMPIS, G. & WURSTER, D. (2024, mai). Using ABM and serious games to create „better AI“. 2024 Annual Modeling and Simulation Conference (ANNSIM) (lk 1–16). IEEE. – doi:10.23919/ANNSIM61499.2024.10732031.

ALLHUTTER, D., CECH, F., FISCHER, F., GRILL, G. & MAGER, A. (2020). Algorithmic profiling of jobseekers in Austria: How austerity politics are made effective. Frontiers in Big Data, 3, 5. – https://doi.org/10.3389/fdata.2020.00005.

ALSTON, P. (2019). Report of the special rapporteur on extreme poverty and human rights. A/74/493. – https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25156.

AMNESTY INTERNATIONAL (13.11.2024). Denmark: Coded injustice: Surveillance and discrimination in Denmark's automated welfare state. – https://www.amnesty.org/en/documents/eur18/8709/2024/en/.

APPELMAN, N., FATHAIGH, R. Ó. & VAN HOBOKEN, J. (2021). Social welfare, risk profiling and fundamental rights: The case of SyRI in The Netherlands. Journal of Intellectual Property, Information Technology and Electronic Commerce Law, 12(4), 257–271.

BERMAN, A., DE FONE LICHT, K. & CARLSSON, V. (2024). Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system. Technology in Society, 76, 102471.

CHIUSI, F., FISCHER, S., KAYSER-BRIL, N. & SPIELKAMP, M. (toim) (2020). Automating society report 2020. Algorith-

mWatch; Bertelsmann Stiftung. – https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/10/Automating-Society-Report-2020.pdf.

DELPIERRE, A., DEMAZIÈRE, D. & EL FATIHI, H. (2024). The stealth legitimization of a controversial policy tool: Statistical profiling in French Public Employment Service. Regulation & Governance, 18(2), 499–512.

DENCIK, L. & KAUN, A. (2020). Datafication and the welfare state. Global Perspectives, 1(1), 12912.

EUBANKS, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St Martin's Press.

JØRGENSEN, R. F. (2021). Data and rights in the digital welfare state: The case of Denmark. Information, Communication & Society, 26(1), 123–138. – https://doi.org/10.1080/1369118X.2021.1934069.

KAUN, A. (2022). Suing the algorithm: The mundanization of automated decision-making in public services through litigation. Information, Communication & Society, 25(14), 2046–2062.

KAUN, A. & MASSO, A. (2025). The data welfare state. Sage Publications Ltd.

KAUN, A., LARSSON, A. O. & MASSO, A. (2024). Automation scenarios: Citizen attitudes towards automated decision-making in the public sector. Information, Communication & Society, 28(7), 1177–1194. – https://doi.org/10.1080/1369118X.2024.2375261.

KÖRTNER, J. & BONOLI, G. (2023). Predictive algorithms in the delivery of public employment services. G. Clegg & N. Durazzi (toim), Handbook of labour market policy in advanced democracies (lk 387–398). Edward Elgar Publishing.

MANN, M. (2020). Technological politics of automated welfare surveillance: Social (and data) justice through critical qualitative inquiry. Global Perspectives, 1(1), 12991.

MISURACA, G. & VAN NOORDT, C. (2020). AI Watch – Artificial intelligence in public services: Overview of the use and impact of AI in public services in the EU (EUR 30255 EN). Publications Office of the European Union. – https://doi.org/10.2760/039619.

MITCHELL, S., POTASH, E., BAROCAS, S., D'AMOUR, A. & LUM, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. Annual Review of Statistics and its Application, 8(1), 141–163.

MÄNNISTE, M., VIHALEMM, T. & TRUMM, A. (2026). Gamifying fairness: Exploring algorithmic decision-making in Estonia's welfare system. P. Ahrweiler & N. Gilbert (toim), Participatory modelling and simulation to improve AI-based public social services (lk 59–72). Springer. – https://link.springer.com/chapter/10.1007/978-3-032-15283-1_5.

OECD (2019). Recommendation of the council on OECD legal instruments artificial intelligence (uuendatud 2024). – https://oecd.ai/en/ai-principles.

PASQUALE, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.

PULTZ, S. (2025). Profiling and subjectification of unemployed people: exploring the case of Slovenian public employment services. Digital Public Employment Services in Action (lk 208–226). Bristol University Press.

RINTA-KAHILA, T., SOMEH, I., GILLESPIE, N., INDULSKA, M. & GREGOR, S. (2024). Managing unintended consequences of algorithmic decision-making: The case of Robodebt. Journal of Information Technology Teaching Cases, 14(1), 165–171. – https://doi.org/10.1177/20438869231165538.

STARKE, C., BALEIS, J., KELLER, B. & MARCINKOWSKI, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. Big Data & Society, 9(2), 20539517221115189. – https://doi.org/10.1177/20539517221115189.

SUN, L. & TANG, Y. (2021). Data-driven discrimination, perceived fairness, and consumer trust – The perspective of consumer attribution. Frontiers in Psychology, 12, 748765. – https://doi.org/10.3389/fpsyg.2021.748765.

VALLI BUTTOW, C. & WEERTS, S. (2022). Public sector information in the European Union policy: The misbalance between economy and individuals. Big Data & Society, 9(2), 20539517221124587. – https://doi.org/10.1177/20539517221124587.

VAN GERVEN, M. & VAN LANCKER, W. (2026). Redefining welfare deservingness: the impact of AI and algorithmic systems on the social investment state. J. Erola, P. Moisio & J. Peltoniemi (toim), Beyond the Nordic Welfare State: Extending Social Investments with Interventions (lk 142–157). Edward Elgar Publishing.

VEALE, M. & BRASS, I. (2019). Administration by algorithm? Public management meets public sector machine learning. K. Yeung & M. Lodge (toim), Algorithmic Regulation (lk 121–149). Oxford University Press. – https://doi.org/10.1093/oso/9780198838494.003.0006.

VIHALEMM, T., MÄNNISTE, M., TRUMM, A. & SOLVAK, M. (2025). Specialists and algorithms: Implementation of AI in the delivery of unemployment services in Estonia. P. Ahrweiler (toim), Participatory artificial intelligence in public social services. Artificial intelligence, simulation and society (lk 97–117). Springer. – https://doi.org/10.1007/978-3-031-71678-2_5.

WEITZ, K., SCHLAGOWSKI, R., ANDRÈ, E., MÄNNISTE, M. & GEORGE, C. (2024, mai). Explaining it your way – findings from a co-creative design workshop on designing XAI applications with AI end-users from the public sector. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (lk 1–14). – https://doi.org/10.1145/3613904.3642563.