

Dirbtinio intelekto reguliavimas viešojoje sveikatos priežiūros sistemoje siekiant užkirsti kelią lyčių ir rasiniam šališkumui

ES pagrindinių teisių chartija grindžiamas medicininio DI reguliavimo modelis

Politikos dokumentas



Bendrai finansuoja
Europos Sąjunga



Ivadas

Dirbtinis intelektas (DI) vis plačiau integruojamas į sveikatos priežiūros sistemas visoje Europoje. Jis naudojamas diagnostikoje, klinikinių sprendimų priėmimo, gydymo planavime, planavimui bei pacientų rizikai prognozuoti. Šios technologijos atveria reikšmingų galimybių didinti sveikatos priežiūros veiksmingumą ir gerinti gydymo rezultatus. Vis dėlto vis daugiau mokslinių įrodymų rodo, kad **DI sistemos gali atkartoti arba sustiprinti jau egzistuojančią nelygybę sveikatos priežiūros srityje**¹.

Yra įvairių priežasčių, dėl kurių DI pagrįsti sprendimai sveikatos priežiūros srityje gali būti šališki. Tarp jų – mokymo duomenų disbalansas, klinikinių kintamųjų matavimo paklaidos, netinkamai parinkti pakaitiniai rodikliai bei neatitikimai tarp aplinkos, kurioje DI sistema buvo sukurta, ir realių jos diegimo sąlygų.² Šališkumas gali atsirasti ir tuomet, kai DI sistemos integruojamos į klinikinius procesus neužtikrinus tinkamo sveikatos priežiūros specialistų pasirengimo arba nesukūrus veiksmingų žmogaus vykdomos priežiūros ir atskaitomybės mechanizmų, leidžiančių stebėti jų veikimą įvairiose pacientų grupėse.³

Empiriniai tyrimai taip pat rodo, kad toks **šališkumas klinikinėje praktikoje gali turėti rimtų pasekmių**. Tyrimų duomenimis, pastebimi reikšmingi skirtumai tarp lyčių bei rasių diagnozuojant ir gydant dažniausiai pasitaikančias ligas. Širdies ir kraujagyslių ligomis sergančioms moterims bei rasinių ir etninių mažumų pacientams diagnozė dažnai nustatoma vėliau, o gydymas ir pacientų priežiūra būna mažiau intensyvi nei vyrams ar baltųjų rasės pacientams.⁴ Rasiniai ir etniniai skirtumai taip pat daro įtaką diabeto priežiūros kokybei ir gydymo rezultatams⁵, o psichikos sveikatos sutrikimai, pavyzdžiui, depresija, tiek moterims, tiek mažumų grupių atstovams diagnozuojami ir gydomi nepakankamai.⁶

¹ Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y., [Dirbtinis intelektas medicinoje: praeitis, dabartis ir ateitis](#), *Stroke and Vascular Neurology* 2(4): 230–243, 2017.

² Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S., [Mašininio mokymosi modelių šališkumo mažinimas medicinoje](#), *Communications Medicine* 1, 25, 2021.

³ Panesar, S., Klot, M., Parrish, R., Fernandez-Miranda, J., Cagle, Y., & Britz, G.W., [Dirbtinio intelekto pažadai ir pavojai neurochirurgijoje](#), *Neurosurgery* 87(1): 33–44, 2020.

⁴ Balla, S., Gomez, S.E., & Rodriguez, F., [Širdies ir kraujagyslių ligų priežiūros ir gydymo rezultaty skirtumai tarp rasinės ir etninės mažumoms priklausančių moterų](#), *Current Treatment Options in Cardiovascular Medicine* 22(12):75, 2020.

⁵ Young, C. & Myers, A. K., [Rasiniai ir etniniai diabeto klinikinės priežiūros bei ligos valdymo netolygumai: naratyvinė apžvalga](#), *Endocrine Practice* 29(4): 295–300, 2023.

⁶ Sorkin, D.H., Ngo-Metzger, Q., Billimek, J., August, K.J., Greenfield, S., & Kaplan, S.H., [Nepakankamai diagnozuojama ir gydoma depresija tarp 2 tipo cukriniu diabetu sergančių pacientų iš įvairių rasinių ir etninių grupių](#), *Diabetes Care* 34(3): 598–600, 2011.

Vienas gerai dokumentuotų pavyzdžių – žurnale „*Journal of Medical Internet Research*“ paskelbtas apžvalginis tyrimas, kuriame nustatyta, kad širdies ligų rizikai prognozuoti naudojami mašininio mokymosi algoritmai skirtingu tikslumu veikia moterų ir vyrų atžvilgiu. Tyrimas parodė, kad moterys buvo sistemingai nepakankamai atstovaujamos mokymo duomenų rinkiniuose, o modeliai moterims dažniau pateikdavo klaidingai neigiamus rezultatus. Tai aiškiai rodo, kad **šališki duomenys gali lemti neteisingus diagnostinius rezultatus**.⁷ panašios tendencijos pastebimos ir rasinio teisingumo aspektu. Kitu tyrimu nustatyta, kad DI modeliai, apmokyti prognozuoti depresijos sunkumą remiantis socialiniuose tinkluose vartojama kalba, gerai veikė baltųjų rasės dalyvių atveju, tačiau gerokai prasčiau – juodaodžių dalyvių. Tai atskleidė rasinius modelių veikimo skirtumus, kuriuos lėmė kalbinės raiškos ypatumai, į kuriuos modeliai neatsižvelgė.⁸

Siekiant valdyti šias su šališkumu susijusias rizikas, **būtinai reguliavimo modeliai, kurie tarptautinius ir Europos etikos bei teisės principus paverstų konkrečiais valdymo mechanizmais**, kuriuos sveikatos priežiūros įstaigos galėtų praktiškai taikyti savo veikloje. Europos Sąjungoje (ES) DI reguliavimo sistema pastaraisiais metais sparčiai keitėsi – priimtas Dirbtinio intelekto aktas (DI aktas), kuriuo nustatomi reikalavimai didelės rizikos DI sistemoms ir jų diegėjams, įskaitant ligonines.⁹ Tačiau šių teisinių reikalavimų perkėlimas į kasdienę sveikatos priežiūros įstaigų praktiką vis dar išlieka dideliu iššūkiu.

Šiame politikos dokumente siūlomas **medicininio DI valdymo reguliavimo modelis**, atitinkantis ES teisyną (acquis), įskaitant ES pagrindinių teisių chartiją, ir yra skirtas viešosioms sveikatos priežiūros įstaigoms bei nevyriausybinėms organizacijoms (PVO). Siūlomas modelis grindžiamas Europos Sąjungos reguliavimo sistema, apimančia DI aktą, Medicinos priemonių reglamentą (MPR), in vitro diagnostikos medicinos priemonių reglamentą (IVPR), Bendrąjį duomenų apsaugos reglamentą (BDAR) ir kuriamą Europos sveikatos duomenų erdvę (ESDE). Juo pagrindinių teisių principai, ypač žmogaus orumo, lygybės, nediskriminavimo ir teisės į sveikatos priežiūrą principai, paverčiami praktinėmis apsaugos priemonėmis, taikytinomis visame DI sistemų gyvavimo cikle. Nuorodos į pasirinktus ne ES valstybių pavyzdžius įtrauktos tik tais atvejais, kai jos suteikia pritaikomų įžvalgų, galinčių papildyti modelio pagrindą sudarančią Europos teisės ir pagrindinių teisių sistemą.

⁷ Straw, I., Rees, G., & Nachev, P., [Mašininio mokymosi algoritimų veikimo skirtumai pagal lytį prognozuojant širdies ligas: apžvalginis tyrimas](#), *Journal of Medical Internet Research* 26: e46936, 2024.

⁸ Rai, S., Stade, E. C., Giorgi, S., Francisco, A., Ungar, L. H., Curtis, B., & Guntuku, S. C., [Pagrindiniai depresiją atspindintys kalbiniai požymiai socialiniuose tinkluose priklauso nuo rasės](#), *Proceedings of the National Academy of Sciences* 121(14): e2319837121, 2024.

⁹ European Union, *Regulation (EU) 2024/1689 of the European Parliament and of the Council, [Suderintos dirbtinio intelekto taisyklės \(DI aktas\)](#)*, PE/24/2024/REV/1, *Official Journal of the European Union*, 13 June 2024.

Be to, rengiant šį reguliavimo modelį remtasi kitose didelės rizikos srityse taikomais reguliavimo principais, įskaitant aviacijos saugos valdymą, finansinių modelių rizikos valdymą, farmakologinį budrumą vaistų reguliavimo srityje ir algoritmų poveikio vertinimą viešajame administravime. Šie metodai suteikia vertingų įžvalgų kuriant valdymo sistemas, galinčias padėti valdyti sudėtingas rizikas, kylančias visame medicininio DI gyvavimo cikle. Šiame modelyje jie pritaikyti taip, kad lyčių ir rasinio teisingumo aspektai būtų aiškiai integruoti į visus DI sistemos gyvavimo ciklo etapus.

Šis modelis sveikatos priežiūros įstaigoms ir pilietinės visuomenės organizacijoms (PVO) siūlo **praktinę valdymo sistemą, apimančią medicininio DI sistemų įsigijimą, validavimą, diegimą, stebėseną ir atskaitomybę**. Integruojant pagrindinių teisių apsaugos priemonės į veiklos procesus, šia sistema siekiama sumažinti su DI grindžiamu sprendimų priėmimu sveikatos priežiūros srityje susijusią lyčių ir rasinio šališkumo riziką bei prisidėti prie teisingesnių sveikatos priežiūros rezultatų visoje Europos Sąjungoje. Be to, dokumente pateikiamas **praktinių politikos priemonių, įrankių ir dokumentų rinkinys, kurį ligoninės ir PVO gali pritaikyti savo nacionaliniam kontekstui**, siekdamos nuosekliai įgyvendinti šį modelį praktikoje.

Europos DI teisinio reguliavimo sistema sveikatos priežiūros srityje

Europos Sąjunga yra sukūrusi daugiasluoksnę medicininio DI teisinio reguliavimo sistemą, kurią sudaro keturi vienas kitą papildantys teisinio reguliavimo lygmenys: Dirbtinio intelekto aktas (DI aktas), medicinos priemonės reglamentuojantys teisės aktai ir susijusios gairės, duomenų apsauga bei sveikatos duomenų valdymas.

Dirbtinio intelekto aktu nustatoma DI sistemų rizika grindžiama reguliavimo sistema. Daugelis sveikatos priežiūros srityje naudojamų DI sprendimų, įskaitant klinikinių sprendimų priėmimo pagalbos priemones ir diagnostikos algoritmus, priskiriami didelės rizikos DI sistemoms, todėl jiems taikomi griežti rizikos valdymo, duomenų valdymo, skaidrumo, žmogaus vykdomos priežiūros ir tikslumo reikalavimai. Taikomi reikalavimai taip pat apima diskriminacinių pasekmių rizikos nustatymą ir mažinimą, atliekant išskaidytą skirtingų gyventojų grupių analizę ir lytį, lytinę tapatybę, etninę kilmę bei kitas socialines kategorijas vertinant kaip horizontaliuosius kintamuosius. **Sveikatos priežiūros įstaigoms, diegiančioms DI sistemas, Dirbtinio intelekto aktas** taip pat nustato konkrečias pareigas. Jos apima pareigą užtikrinti, kad DI sistemos būtų naudojamos laikantis teikėjo instrukcijų, stebėti jų veikimą, tvarkyti veiklos žurnalus ir pranešti kompetentingoms institucijoms apie rimtus incidentus, inter alia tais atvejais, kai tokie incidentai gali turėti neproporcingą poveikį skirtingoms pacientų grupėms.

Kalbant apie Dirbtinio intelekto akto taikymą, 2025 m. jau įsigaliojo draudžiamos praktikos nuostatos ir pradėtos taikyti kai kurios priežiūros bei valdymo taisyklės.¹⁰ Kitas svarbus etapas – 2026 m. rugpjūtis, kai bus pradėta taikyti didžioji dalis Akto reikalavimų, įskaitant III priede nurodytoms didelės rizikos DI sistemoms taikomas nuostatas ir 50 straipsnyje nustatytus skaidrumo reikalavimus. Reglamentuojamuose gaminiuose integruotoms didelės rizikos DI sistemoms (įskaitant daugelį DI sistemų, naudojamų medicinos priemonėse) Akto reikalavimai bus pradėti taikyti nuo 2027 m. rugpjūčio.

Dirbtinio intelekto akto nuostatos, svarbiausios ligoninėms praktiniu atžvilgiu:

- **Didelės rizikos DI sistemoms taikomi reikalavimai.** Teikėjai privalo įdiegti rizikos valdymo sistemą, apimančią visą DI sistemos gyvavimo ciklą (9 straipsnis), ir taikyti duomenų valdymo praktiką, kuri apimtų: aiškius šališkumo nustatymo ir mažinimo reikalavimus (10 straipsnis), įskaitant duomenų reprezentatyvumo ir galimų DI sistemos veikimo skirtumų skirtingose gyventojų grupėse vertinimą, atliekant detalizuotą analizę pagal lytį, socialinę lytį, rasinę ar etninę kilmę; skaidrumo užtikrinimo priemonės ir informacijos teikimą diegėjams (13 straipsnis); žmogaus vykdomą priežiūrą, skirtą automatizavimo šališkumui išvengti ir sudaryti galimybę nepaisyti DI sistemos sprendimo arba sustabdyti jos veikimą (14 straipsnis); taip pat tikslumo, patikimumo ir kibernetinio saugumo kontrolės priemonės, ypatingą dėmesį skiriant grįžtamojo ryšio ciklams ir DI sistemoms būdingoms kibernetinėms atakoms (15 straipsnis), kurios, jei nebūtų tinkamai valdomos, galėtų dar labiau sustiprinti esamą nelygybę.
- **DI sistemų diegėjams taikomi reikalavimai.** Ligoninės ir kiti DI sistemų diegėjai privalo naudoti didelės rizikos DI sistemas laikydamiesi teikėjo instrukcijų (26 straipsnis); paskirti kompetentingus asmenis, atsakingus už žmogaus vykdomą priežiūrą (26 straipsnis); užtikrinti, kad įvesties duomenys būtų reprezentatyvūs ir išsamūs tais atvejais, kai diegėjas kontroliuoja įvedamus duomenis (26 straipsnis), taip pat, kai tai įmanoma, atsižvelgti į klinikinių populiacijų demografinę įvairovę; stebėti DI sistemos veikimą (26 straipsnis); ne trumpiau kaip šešis mėnesius saugoti automatiškai sugeneruotus veiklos žurnalus (26 straipsnis); ir laiku pranešti teikėjams bei kompetentingoms institucijoms apie rimtus incidentus (26 ir 73 straipsniai), ypač tais atvejais, kai jie gali lemti nevienodus klinikinius rezultatus skirtingoms pacientų grupėms.

¹⁰ European Commission, [Dirbtinio intelekto akto įgyvendinimo planas](#), Brussels: European Commission, 2024.

- **Poveikio pagrindinems teisėms vertinimas.** Prieš diegdamos tam tikras didelės rizikos DI sistemas, viešosios teisės reglamentuojamos įstaigos (taip pat viešąsias paslaugas teikiantys privatūs subjektai) privalo atlikti poveikio pagrindinems teisėms vertinimą, kuriame aprašomi numatomi procesai, paveiktos grupės, galimos rizikos, žmogaus vykdomos priežiūros priemonės, rizikos mažinimo ir skundų nagrinėjimo mechanizmai, o vėliau apie tai informuoti rinkos priežiūros instituciją (28 straipsnis). Šis reikalavimas ypač svarbus daugelio valstybių narių viešosioms ligoninėms, nes sudaro galimybę dar prieš diegiant DI sistemą nustatyti diskriminacijos riziką ir neteisingo poveikio skirtingoms gyventojų grupėms riziką bei imtis priemonių jai mažinti, remiantis detalizuota analize pagal lytį, socialinę lytį, rasinę ir etninę kilmę.

Be Dirbtinio intelekto akto, medicinos srityje naudojamoms **DI sistemoms taip pat taikomi medicinos priemonės reglamentuojantys teisės aktai ir susijusios gairės**, įskaitant Reglamentą (ES) 2017/745 dėl *medicinos priemonių* (MPR)¹¹ ir Reglamentą (ES) 2017/746 dėl *in vitro diagnostikos medicinos priemonių* (IVPR)¹², kuriais reglamentuojama medicinos technologijų sauga ir veiksmingumas. Šiuose reglamentuose nustatyta, kad medicinos priemonių programinei įrangai, įskaitant DI grindžiamas priemones, turi būti atliekamas klinikinis vertinimas ir stebėseną po pateikimo rinkai. Kai tai leidžia turimi duomenys, taip pat turi būti vertinama, ar tokios priemonės vienodai veiksmingai veikia skirtingose pacientų grupėse, remiantis detalizuota jų analize. Be to, Medicinos priemonių koordinavimo grupės (MDCG) gairėse dėl *programinės įrangos kvalifikavimo ir klasifikavimo pagal MPR ir IVPR* pateikiamos rekomendacijos, kaip kvalifikuoti ir klasifikuoti programinę įrangą pagal šiuos reglamentus.¹³ Jose pabrėžiama, kad programinės įrangos klasifikavimas priklauso nuo jos numatytos paskirties, o ne nuo diegimo būdo (pavyzdžiui, ar ji veikia debesijoje, ar yra integruota į medicinos priemonę). Gairėse taip pat pateikiami ligoninių informacinių sistemų ir klinikinių sprendimų priėmimo pagalbos programinės įrangos pavyzdžiai, pabrėžiant būtinybę įvertinti galimą skirtingą poveikį įvairioms pacientų grupėms.¹⁴

Dirbtinio intelekto akto ir galiojančių medicinos priemonės reglamentuojančių teisės aktų tarpusavio santykis paaiškinamas *Gairėse dėl Dirbtinio intelekto akto ir MPR bei*

¹¹ European Union, [Reglamentas \(ES\) 2017/745 dėl medicinos priemonių \(MPR\)](#), Official Journal of the European Union, L117, Brussels: EU, 2017.

¹² European Union, [Reglamentas \(ES\) 2017/746 dėl in vitro diagnostikos medicinos priemonių \(IVPR\)](#), Official Journal of the European Union, L117, Brussels: EU, 2017.

¹³ Medical Device Coordination Group, Programinės įrangos kvalifikavimo ir klasifikavimo pagal [Reglamentą \(ES\) 2017/745 dėl medicinos priemonių \(MPR\) ir Reglamentą \(ES\) 2017/746 dėl in vitro diagnostikos medicinos priemonių \(IVPR\) gairės](#), Brussels: MDCG, 2019.

¹⁴ Medical Device Coordination Group, [Programinės įrangos, naudojamos kaip medicinos priemonė, klinikinio vertinimo \(MPR\) ir veikimo vertinimo \(IVPR\) gairės](#), Brussels: MDCG, 2020.

IVPR sąveikos.¹⁵ Dokumente aiškinama, kad Dirbtinio intelekto aktas taikomas kartu su dviem minėtais reglamentais. Praktikoje tai reiškia, kad daugeliui **medicininio DI sistemų vienu metu taikomi abiejų reguliavimo sistemų reikalavimai**. Gairėse taip pat paaiškinama, kada medicininis DI pagal DI aktą laikomas didelės rizikos DI sistema. Paprastai taip yra tais atvejais, kai pati DI sistema yra medicinos priemonė (arba jos saugos komponentas), todėl jai taikoma paskelbtosios įstaigos atliekama trečiosios šalies atitikties vertinimo procedūra. Ligoninėms, įsigyjantioms DI sistemas, tai turi svarbių praktinių pasekmių. Jos turėtų reikalauti, kad tiekėjai pagrįstų savo sistemų atitiktį taikomiems reguliavimo reikalavimams, įskaitant jų klasifikavimą pagal MPR arba IVPR, paskelbtosios įstaigos dalyvavimą tais atvejais, kai to reikalaujama, ir atitiktį Dirbtinio intelekto akto didelės rizikos DI sistemoms taikomiems reikalavimams. Taip pat turėtų būti pateikiami duomenys apie DI sistemos veikimą skirtingose gyventojų grupėse, remiantis detalizuota analize, ir, kai aktualu, taikomas šališkumo mažinimo priemonės. DI sistemų įsigijimo sutartyse taip pat turėtų būti numatyta tiekėjų pareiga pateikti dokumentaciją bei informaciją apie skaidrumą ir žmogaus vykdomos priežiūros priemones, kurių DI sistemų diegėjams reikia pagal Dirbtinio intelekto aktą, įskaitant nuostatas, skirtas užtikrinti teisingus rezultatus visiems pacientams.

Praktikoje medicininio DI valdymui **svarbūs ir keli kiti Europos Sąjungos teisės aktai**. Vienas svarbiausių – *Bendrasis duomenų apsaugos reglamentas* (BDAR), kuris medicininio DI srityje yra itin reikšmingas, nes sveikatos duomenys priskiriami specialių kategorijų asmens duomenims, o diegiant DI sistemas ligoninėse dažnai privaloma atlikti *poveikio duomenų apsaugai vertinimą*. Dirbtinio intelekto akte (26 straipsnio 9 dalis ir 27 straipsnio 4 dalis) aiškiai susiejamos DI sistemų diegėjų pareigos su reikalavimu atlikti poveikio duomenų apsaugai vertinimą. Jame taip pat numatyta, kad teikėjų pateikta informacija turi būti naudojama atliekant šį vertinimą, o poveikio pagrindinėms teisėms vertinimas laikomas jį papildančia priemone. Atliekant abu vertinimus gali būti analizuojamas galimas nevienodas poveikis skirtingoms gyventojų grupėms, išskirtoms pagal lytį, socialinę lytį, rasinę ar etninę kilmę. Be to, *Europos sveikatos duomenų erdvės* (ESDE) tikslas – sudaryti sąlygas saugiai naudotis sveikatos duomenimis sveikatos priežiūros paslaugų teikimo ir mokslinių tyrimų tikslais, kartu užtikrinant aukštą duomenų apsaugos lygį.¹⁶ Tikimasi, kad ESDE atliks svarbų vaidmenį kuriant reprezentatyvesnius medicininio DI mokymo duomenų rinkinius, sudarys prielaidas vertinti DI sistemas teisingumo požiūriu ir padės mažinti riziką, kad tam tikrų demografinių grupių pacientai bus nepakankamai atstovaujami duomenų rinkiniuose.¹⁷

¹⁵ Medical Device Coordination Group & European AI Board, [Gairės dėl MPR, IVPR ir Dirbtinio intelekto akto sąveikos](#), Brussels: European Commission, 2025.

¹⁶ European Union, [Reglamentas dėl Europos sveikatos duomenų erdvės \(ESDE\)](#), Brussels: European Union, 2024.

¹⁷ European Commission, [Europos sveikatos duomenų erdvė \(EDSE\)](#), Brussels: European Commission, 2025.

Galiausiai, kalbant apie kibernetinį saugumą, Europos Sąjungos kibernetinio saugumo agentūra (ENISA) yra parengusi gaires dėl DI sistemų kibernetinio saugumo praktikos. Gairėse pažymima, kad esami organizaciniai standartai (pavyzdžiui, ISO saugumo ir kokybės valdymo standartai) gali būti pritaikomi DI sistemoms.¹⁸ MDCG *gairėse dėl medicinos priemonių kibernetinio saugumo* taip pat nurodoma, kaip gamintojai turėtų užtikrinti MPR ir IVPR nustatytų su kibernetiniu saugumu susijusių reikalavimų laikymąsi, taikydami rizikos valdymo priemonės ir užtikrindami minimalius kibernetinio saugumo standartus.¹⁹ Svarbu pabrėžti, kad veiksmingas kibernetinio saugumo valdymas yra viena iš teisingo DI diegimo prielaidų – duomenų vientisumo, prieigos kontrolės ir sistemų patikimumo užtikrinimas padeda išvengti netyčinio lyčių ir rasinės ar etninės nelygybės didėjimo

Visi šie teisės aktai kartu sudaro sudėtingą teisinio reguliavimo sistemą, kurioje veikia sveikatos priežiūros srityje naudojamos DI sistemos. Šiame dokumente siūlomas Europos viešosioms ligoninėms skirtas DI reguliavimo modelis parengtas remiantis keturiais pagrindiniais teisinio reguliavimo lygmenimis, kurie sudaro medicininio DI reguliavimo pagrindą. Modelyje numatyti mechanizmai, skirti viso DI sistemos gyvavimo ciklo metu stebėti ir mažinti nevienodo poveikio riziką skirtingoms gyventojų grupėms, išskirtoms pagal lytį, socialinę lytį, rasinę ar etninę kilmę.

Europos ir tarptautinės atsakingo DI valdymo sistemos sveikatos priežiūros srityje

Svarbu pažymėti, kad pirmiau aptartos Europos reguliavimo ir standartizacijos sistemos sudaro pagrindinį medicininio DI valdymo teisinį pagrindą Europos Sąjungoje. Be ES medicininio DI teisinio reguliavimo sistemos, DI rizikos valdymo gairės taip pat nustato keli teisiškai neprivalomi tarptautiniai dokumentai. Šiame **skyriuje aptariamos papildomos valdymo sistemos pateikiamos siekiant parodyti, kokie jų įgyvendinimo sprendimai galėtų būti pritaikyti ES kontekste, kartu išlaikant visišką jų atitiktį Europos teisės reikalavimams ir pagrindinių teisių principams**. Šie pavyzdžiai pasitelkiami kaip kitose jurisdikcijose taikomų sprendimų analogai, kuriais iliustruojami praktikoje jau įdiegti panašūs visą DI sistemos gyvavimo ciklą apimantys valdymo mechanizmai, ypač modelių rizikos valdymo, algoritmų audito ir institucinio poveikio vertinimo srityse.

Pavyzdžiui, Nacionalinio standartų ir technologijų instituto (NIST) parengta *DI rizikos valdymo sistema (angl. AI Risk Management Framework, RMF)* siūlo išsamų požiūrį į DI

¹⁸ European Union Agency for Cybersecurity (ENISA), [Dirbtinio intelekto kibernetinis saugumas ir standartizacija](#), Athens: ENISA, 2023.

¹⁹ Medical Device Coordination Group, [MDCG 2019-16 1 red. – Medicinos priemonių kibernetinio saugumo gairės](#), Brussels: European Commission, 2020.

rizikos valdymą viso DI sistemos gyvavimo ciklo metu.²⁰ Ši sistema grindžiama keturiais pagrindiniais elementais: valdymu, rizikų nustatymu, rizikos vertinimu ir rizikos valdymu. Joje teisingumas ir žalingo šališkumo prevencija aiškiai įvardijami kaip esminiai patikimo DI požymiai. Be to, nustatant ir vertinant riziką organizacijos skatinamos identifikuoti ir įvertinti galimą nevienodą poveikį skirtingoms pacientų grupėms, atsižvelgiant į lytį, socialinę lytį, rasinę ir etninę kilmę. Šios sistemos tikslas – užtikrinti, kad patikimumo ir rizikos kontrolės priemonės būtų integruotos visuose DI sistemos kūrimo, diegimo ir stebėsenos etapuose, derinant techninį vertinimą su organizacine atskaitomybe. Ligoninėms ši DI rizikos valdymo sistema ypač vertinga tuo, kad jos principus galima lengvai pritaikyti praktikoje, pavyzdžiui, **aiškiai apibrėžiant už teisingumo priežiūrą atsakingų asmenų funkcijas, sistemingai vertinant DI sistemų veikimą skirtingose gyventojų grupėse, remiantis detalizuota analize, ir taikant dokumentais pagrįstas rizikos mažinimo strategijas**, kartu išlaikant platesnį pagrindinių teisių apsaugos požiūrį.

Valdymo sistemų srityje svarbus vaidmuo tenka Tarptautinės standartizacijos organizacijos standartui *ISO/IEC 42001*, kuriame nustatyti organizacinės DI valdymo sistemos reikalavimai.²¹ Šiame standarte daugiausia dėmesio skiriama atsakingą DI užtikrinančiai politikai, tikslams, procesams ir nuolatiniam tobulinimui. Nors jame išsamiai nereglamentuojamos konkrečios žalos rūšys, standartas skatina į organizacijos politiką ir rizikos kontrolės priemones integruoti teisingumo ir nediskriminavimo principus, taip sudarant sąlygas nuosekliai valdyti galimą lyčių ir rasinį šališkumą. Šį standartą papildė *ISO/IEC 23894*, kuriame pateikiamos gairės, kaip rizikos valdymą integruoti į DI sistemų kūrimo ir diegimo procesus, įskaitant rizikų, susijusių su netyčiniu ar neproporcingu poveikiu skirtingoms gyventojų grupėms, nustatymą ir valdymą.²² Medicininio DI srityje ypač svarbus ir standartas *ISO 14971*, kuriame nustatyti medicinos priemonių reguliavime plačiai taikomi rizikos valdymo principai, apimantys sistemingą pavojų nustatymą, rizikos vertinimą, rizikos kontrolę ir stebėseną viso gyvavimo ciklo metu.²³ Šie procesai gali būti taikomi ir vertinant nevienodą DI sistemų veikimą ar rezultatus skirtingose gyventojų grupėse, remiantis detalizuota analize, kaip galimos žalos nustatymo dalį. Aptarti standartai vertingi tuo, kad siūlo **aiškią valdymo sistemos struktūrą, kurią gali perimti ligoninės**. Jie taip pat sukuria bendrą terminologinį pagrindą, leidžiantį DI sistemų įsigijimo ir sutarčių su tiekėjais sudarymo etapuose aiškiai apibrėžti su teisingumu susijusius reikalavimus,

²⁰ National Institute of Standards and Technology, [Dirbtinio intelekto rizikos valdymo sistema \(AI RMF 1.0\)](#), Gaithersburg, MD: NIST, 2023.

²¹ International Organization for Standardization, [ISO/IEC 42001:2023 – Dirbtinio intelekto valdymo sistema](#), Geneva: ISO, 2023.

²² International Organization for Standardization, [ISO/IEC 23894:2023 – Dirbtinio intelekto rizikos valdymas](#), Geneva: ISO, 2023.

²³ International Organization for Standardization, [ISO 14971:2019 – Rizikos valdymo taikymas medicinos priemonėms](#), Geneva: ISO, 2019.

pavyzdžiui, DI sistemų veikimo vertinimą skirtingose gyventojų grupėse, remiantis detalizuota analize, ir šališkumo mažinimo priemonės.

Svarbių gairių teikia ir etikos principais grindžiamos DI valdymo sistemos. Europos Komisijos parengtos *Patikimo dirbtinio intelekto etikos gairės*, nors ir nėra teisiškai privalomos, yra reikšmingas politikos formavimo dokumentas.²⁴ Jose nustatyti septyni pagrindiniai atsakingo DI reikalavimai, tarp jų – įvairovė, nediskriminavimas ir teisingumas. Gairėse aiškiai pabrėžiama būtinybė vengti šališkumo ir vertinti DI sistemų poveikį skirtingoms gyventojų grupėms, atsižvelgiant į tokius demografinius požymius kaip lytis, socialinė lytis, rasinė ir etninė kilmė. Šias gaires praktiškai įgyvendina *Patikimo dirbtinio intelekto vertinimo klausimynas (ALTAI)* – savanoriška įsivertinimo priemonė, parengta Europos Komisijos Aukšto lygio ekspertų grupės dirbtinio intelekto klausimais.²⁵ Sveikatos priežiūros reguliavimo kontekste ALTAI ypač vertingas dėl **klausimynu grindžiamos struktūros, kurią galima pritaikyti rengiant praktinius DI sistemų įsigijimo ir diegimo kontrolinius sąrašus**. Jie padeda suinteresuotosioms šalims įvertinti tokius aspektus kaip mokymo duomenų reprezentatyvumas, nevienodų klinikinių rezultatų rizika ar apsaugos priemonių pakankamumas siekiant išvengti diskriminacinio poveikio. Taip teisingumo principai integruojami į kasdienį gydytojų, informacinių technologijų specialistų ir ligoninių vadovų sprendimų priėmimą.

Žvelgiant iš žmogaus teisių perspektyvos, Europos Tarybos *Pagrindų konvencija dėl dirbtinio intelekto* yra svarbus tarptautinis teisinis instrumentas, aiškiai grindžiamas žmogaus orumo, skaidrumo, atskaitomybės, lygybės, nediskriminavimo ir privatumo principais. Ji suteikia teisinį pagrindą spręsti diskriminacinio poveikio problemas, įskaitant tas, kurios kyla dėl lyčių ar rasinio šališkumo DI sistemose.²⁶ Panašiai Ekonominio bendradarbiavimo ir plėtros organizacijos (EBPO) *Dirbtinio intelekto principai*²⁷ bei UNESCO *Rekomendacija dėl dirbtinio intelekto etikos*²⁸ pateikia tarptautines norminio pobūdžio gaires, skatinančias nustatyti ir mažinti nevienodą ar neteisingą DI sistemų poveikį skirtingoms gyventojų grupėms, remiantis detalizuota analize. Taip pabrėžiama, kad DI sistemos neturėtų atkartoti ar stiprinti struktūrinės nelygybės. Sveikatos priežiūros srityje Pasaulio sveikatos organizacija yra suformulavusi šešis pagrindinius DI etikos principus: autonomiją, gerovę, paaiškinamumą,

²⁴ European Commission, [Patikimo dirbtinio intelekto etikos gairės](#), Brussels: European Commission, 2019.

²⁵ European Commission, [Patikimo dirbtinio intelekto vertinimo kontrolinis sąrašas \(ALTAI\): įsivertinimo priemonė](#), Brussels: European Commission, 2020.

²⁶ Council of Europe, [Pagrindų konvencija dėl dirbtinio intelekto \(CAI\)](#), Strasbourg: Council of Europe, 2023.

²⁷ Organisation for Economic Co-operation and Development, [EBPO dirbtinio intelekto principai](#), Paris: OECD, 2019.

²⁸ UNESCO, [Rekomendacija dėl dirbtinio intelekto etikos](#), Paris: UNESCO, 2021.

atsakomybę, teisingumą ir jautrumą poreikiams.²⁹ **Teisingumo principas aiškiai įpareigoja vertinti prieigos prie sveikatos priežiūros, DI sistemų veikimo ir jų rezultatų skirtumus įvairiose pacientų grupėse, atsižvelgiant į lytį, socialinę lytį, rasinę ir etninę kilmę.**

Visas šias tarptautines DI valdymo sistemas vienija tai, kad jos tiesiogiai atitinka ES pagrindinių teisių chartijoje įtvirtintus Žmogaus orumo, lygybės, nediskriminavimo ir teisės į sveikatos priežiūrą principus. Jos taip pat pabrėžia visą DI sistemos gyvavimo ciklą apimančią valdymą, kurio principai gali būti tiesiogiai pritaikyti kuriant praktines kontrolės priemones sveikatos priežiūros įstaigose.

Tarpsektorinių reguliavimo modelių taikymo pamokos

Tokiuose didelės rizikos sektoriuose kaip aviacija, bankininkystė ir farmacija jau seniai taikomos pažangios valdymo sistemos, skirtos technologinėms rizikoms valdyti. Šiuose sektoriuose sukaupta reguliavimo patirtis gali būti sėkmingai pritaikyta medicininio DI valdymui sveikatos priežiūros srityje ir padėti spręsti su lyčių bei rasiniu šališkumu susijusius iššūkius.

Nors lyčių ir rasinis šališkumas siaurąja prasme nėra technologinės rizikos, jie tampa sociotechninėmis rizikomis, kai duomenys, algoritmai ir organizacinė praktika sąveikauja su struktūrine nelygybe sveikatos priežiūros sistemoje. Šiame kontekste sąvoka „rizikos šaltinis“ reiškia ne pačią šališkumo priežastį, o jo pasireiškimą sistemos lygmeniu – DI sistemų rezultatuose, priimamuose sprendimuose ir klinikiniuose rezultatuose. Sveikatos priežiūros srityje jau sukaupta daug įrodymų apie tokį poveikį, įskaitant nustatytus diagnozavimo, gydymo ir rizikos prognozavimo skirtumus tarp gyventojų grupių, išskirtų pagal lytį, socialinę lytį, rasinę ar etninę kilmę. Todėl kitų sektorių pavyzdžiai papildo sveikatos priežiūros srityje sukauptus įrodymus, parodydami, kaip praktiškai gali būti taikomi pažangūs sudėtingų ir didelį poveikį darančių sistemų valdymo mechanizmai, o sveikatos priežiūros srityje atlikti empiriniai tyrimai pagrindžia tokių reguliavimo priemonių būtinybę.

Ypač reikšmingas pavyzdys – aviacijoje taikoma *saugos valdymo sistema* (angl. *Safety Management System, SMS*). Tarptautiniuose civilinės aviacijos teisės aktuose sauga suprantama ne kaip vienkartinis sertifikavimo etapas, o kaip nuolatinis visos organizacijos mastu vykdomas procesas, leidžiantis priimti sprendimus remiantis rizikos šaltinių nustatymu, rizikos valdymu, saugos užtikrinimu (stebėseną ir auditu), pokyčių valdymu ir nuolatiniu tobulinimu.³⁰ Pritaikius šį požiūrį medicininiam DI, būtų galima sukurti „DI saugos valdymo sistemą“, kurioje aiškiai integruoti teisingumo principai.

²⁹ World Health Organization, [Dirbtinio intelekto sveikatos priežiūroje etika ir valdymas](#), Geneva: WHO, 2021.

³⁰ International Civil Aviation Organization, [Saugos valdymo vadovas \(Doc 9859\)](#), 4th ed., Montréal: ICAO, 2018.

Tokia sistema galėtų apimti: (1) rizikos šaltinių ir beveik įvykusių incidentų registravimo kanalą, skirtą sveikatos priežiūros specialistams pranešti apie galimus lyčių ar rasinio šališkumo atvejus; (2) formalų kiekvienos DI sistemos rizikų registrą, kuriame būtų dokumentuojami DI sistemos veikimo skirtumai skirtingose pacientų grupėse, remiantis detalizuota analize; (3) pokyčių valdymo procedūras, leidžiančias įvertinti modelių atnaujinimų, duomenų pokyčių ar klinikinių procesų pasikeitimų poveikį teisingumui; ir (4) periodines saugos užtikrinimo peržiūras, kartu numatant korekcinis veiksmus techninėms rizikoms ir nustatyti lyčių ar rasinei nelygybei šalinti. Toks požiūris glaudžiai dera su Dirbtinio intelekto akto 9 ir 72 straipsniuose įtvirtintu reikalavimu taikyti visą DI sistemos gyvavimo ciklą apimantį rizikos valdymą ir vykdyti stebėseną po pateikimo rinkai.

Kitas naudingas pavyzdys – finansų sektoriaus reguliavimas, kuriame sudėtingi kiekybiniai įrankiai laikomi modeliais, kurių kūrimui, nepriklausomam validavimui ir priežiūrai taikomi griežti reikalavimai. Panašios modelių rizikos valdymo sistemos taikomos, pavyzdžiui, Jungtinių Amerikos Valstijų bankų priežiūros praktikoje (pvz., *JAV Federalinio rezervo sistemos gairėse dėl modelių rizikos valdymo*) ir parodo, kaip praktiškai galima organizuoti modelių validavimą, priežiūrą ir nepriklausomą vertinimą.³¹ Tokį požiūrį būtų galima taikyti ir medicininio DI sistemoms, siekiant užtikrinti, kad **klinikinių sprendimų priėmimo pagalbos priemonės prieš jas pradedant naudoti sveikatos priežiūros įstaigose būtų nepriklausomai validuojamos** (vidaus arba išorės ekspertų), kartu **aiškiai įvertinant jų veikimą skirtingose gyventojų grupėse pagal lytį, socialinę lytį, rasinę ar etninę kilmę**. Atliekant validavimą turėtų būti dokumentuojama DI sistemos numatyta paskirtis, taikymo ribos, žinomi veikimo sutrikimų scenarijai ir veikimo rezultatai skirtingose gyventojų grupėse, remiantis detalizuota analize, taip pat nustatyti galimi skirtumai ir taikomos jų mažinimo priemonės. Pakartotinis validavimas turėtų būti atliekamas kiekvieną kartą, kai DI sistema diegiama naujoje įstaigoje, pradedama taikyti kitai pacientų grupei, integruojama su nauja įranga ar klinikiniais protokolais, atnaujinama teikėjo arba pastebima jos veikimo kokybės blogėjimo požymių. Toks požiūris atitinka Dirbtinio intelekto akto 13, 14 ir 60 straipsniuose nustatytus reikalavimus dėl DI sistemų testavimo, informacijos teikimo diegėjams ir žmogaus vykdomos priežiūros, skirtos automatizavimo šališkumui nustatyti ir jam užkirsti kelią, taip užtikrinant, kad DI grindžiama sveikatos priežiūra būtų saugi ir teisinga visoms gyventojų grupėms.

Kalbant apie DI sistemų priežiūrą po jų įdiegimo, geru pavyzdžiu gali būti farmacijos sektoriaus reguliavimas. Farmakologinio budrumo sistemoje nepageidaujamos reakcijos ir saugos signalai sistemingai renkami ir analizuojami siekiant nustatyti naujai kylančias rizikas. Visų pirma, Europos vaistų agentūros *Geros farmakologinio budrumo praktikos (GVP) VI modulyje* išsamiai apibrėžiami reikalavimai, taikomi įtariamų

³¹ Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, [Priežiūros gairės dėl modelių rizikos valdymo \(SR 11-7\)](#), Washington, D.C., 4 April 2011.

nepageidaujamų reakcijų pranešimų rinkimui, tvarkymui ir teikimui, taip sukuriant nuoseklią incidentų registravimo ir tolesnių veiksmų sistemą.³² Panašiu principu galėtų būti sukurta ir medicininio DI priežiūros sistema („**DI budrumo**“ sistema), kuri leistų ligoninėms registruoti ir analizuoti incidentus, susijusius su DI grindžiamais klinikiniais sprendimais, ypatingą dėmesį skiriant jų poveikiui teisingumui. Tokia sistema galėtų apimti: (1) sistemingą incidentų registravimą, leidžiantį nustatyti sveikatos priežiūros ar gydymo rezultatų skirtumus, susijusius su lytimi, socialine lytimi, rasine ar etnine kilmė; (2) privalomą rimtų incidentų, kuriems įtakos galėjo turėti DI sistemų šališkumas, dokumentavimą; (3) pareigą pranešti teikėjams ir kompetentingoms institucijoms apie su teisingumu susijusias problemas tais atvejais, kai šališkumas galėjo lemti didelę žalą, pavyzdžiui, klaidingą diagnozę ar nevienodą pacientų gydymą, o mažesnės reikšmės atvejus toliau stebėti vykdant vidaus kokybės priežiūrą; ir (4) tendencijų bei saugos signalų nustatymo funkcijas, leidžiančias identifikuoti naujai atsirandančius šališkumo modelius skirtinguose skyriuose, gydymo įstaigose ar platesniame sveikatos priežiūros tinkle. Tokia stebėsenos sistema sudarytų sąlygas laiku taikyti prevencines rizikos mažinimo priemones ir nuolat tobulinti DI sistemų naudojimą, užtikrinant, kad DI grindžiama sveikatos priežiūra būtų ne tik saugi, bet ir teisinga.

Galiausiai, viešajame sektoriuje taikomos valdymo priemonės, tokios kaip algoritmų poveikio vertinimas ir šališkumo auditai, suteikia praktinių galimybių įvertinti automatizuotų sprendimų priėmimo sistemų poveikį visuomenei, įskaitant galimus skirtumus tarp gyventojų grupių pagal lytį, socialinę lytį, rasinę ar etninę kilmę. Šios priemonės leidžia etikos principus ir žmogaus teisių standartus paversti praktiniais įrankiais, padedančiais organizacijoms dar prieš diegiant DI sistemas numatyti galimą nelygybę ir imtis priemonių jai mažinti. Pavyzdžiui, Kanados federalinė *Algoritmų poveikio vertinimo sistema* įvertina DI sistemų keliamą riziką ir susieja ją su konkrečiomis rizikos mažinimo priemonėmis.³³ Ši priemonė aiškiai parodo, kad DI sistemų diegimas turi būti vertinamas pirmiausia atsižvelgiant į jų galimą poveikį. Be to, Niujorko miesto taikomas automatizuotų darbuotojų atrankos priemonių reguliavimo modelis numato privalomus šališkumo auditus ir pareigą viešai atskleisti informaciją apie tokių sistemų naudojimą.³⁴ Panašų „**DI poveikio vertinimo ir audito**“ priemonių rinkinį galėtų taikyti ir ligoninės bei sveikatos priežiūros srityje veikiančios pilietinės visuomenės organizacijos. Jų galėtų sudaryti **struktūruotas į teisingumą orientuotas klausimynas ir periodiškai atliekami nepriklausomi teisingumo auditai**, kurių rezultatai būtų apibendrinami viešai skelbiamoje ataskaitoje apie DI sistemų veikimą skirtingose gyventojų grupėse, remiantis detalizuota analize, bei taikomas

³² European Medicines Agency, [Geros farmakologinio budrumo praktikos \(GVP\) VI modulis. Įtariamų nepageidaujamų reakcijų rinkimas ir pranešimas apie jas](#), London: EMA, 2017

³³ Treasury Board of Canada Secretariat, [Algoritmų poveikio vertinimas](#), Ottawa: Government of Canada, 2026.

³⁴ New York City Department of Consumer and Worker Protection, [Automatizuotos sprendimų priėmimo priemonės įdarbinimo srityje](#), New York: City of New York, 2023.

rizikos mažinimo priemonės. Tokios priemonės prisidėtų prie didesnio skaidrumo, viešosios atskaitomybės ir pacientų pasitikėjimo DI grindžiama sveikatos priežiūra.

Į šį modelį įtraukti reguliavimo pavyzdžiai iš aviacijos saugos, finansinių modelių rizikos valdymo, farmakologinio budrumo ir algoritmų valdymo sričių pasirinkti sąmoningai, siekiant perkelti jose taikomus ir pasiteisinusius valdymo principus. Šios sritys pasirinktos todėl, kad jose praktiškai taikomi formalizuoti rizikos nustatymo, stebėsenos ir mažinimo metodai, skirti sudėtingoms sistemoms, kuriose gali pasireikšti nevienodas poveikis skirtingiems asmenims. Kai kuriose iš jų jau taikomos ir su teisingumu susijusios priemonės, pavyzdžiui, demografinių rizikos veiksnių stratifikavimas farmakologinio budrumo srityje, modelių validavimas ir šališkumo kontrolė finansų sektoriaus priežiūroje, taip pat sistemingas nepageidaujamų įvykių, turinčių poveikį konkrečioms gyventojų grupėms, registravimas. Šių sričių svarba grindžiama ne tik jų konceptuali panašumu į medicininio DI reguliavimą, bet ir tuo, kad jose sėkmingai taikomi visą sistemos gyvavimo ciklą apimantys valdymo principai, privalomas incidentų registravimas ir nepriklausomos priežiūros mechanizmai, kuriuos galima tiesiogiai pritaikyti valdant su lyčių ir rasiniu teisingumu susijusias rizikas medicininio DI sistemose.

ES pagrindinių teisių chartija grindžiamo DI reguliavimo modelio viešosioms ligoninėms ir pilietinės visuomenės organizacijoms gairės

Šiame skyriuje pristatoma DI reguliavimo modelio struktūra, **grindžiama ES pagrindinių teisių chartijos principais ir paverčianti juos praktinėmis procedūromis, kurias savo veikloje gali taikyti ligoninės ir pilietinės visuomenės organizacijos**. Chartija sudaro norminį sveikatos technologijų valdymo Europoje pagrindą, visų pirma įtvirtindama žmogaus orumo, lygybės, nediskriminavimo ir teisės į sveikatos priežiūrą principus.³⁵ Šios teisės ypač svarbios tais atvejais, kai automatizuotos sistemos daro įtaką klinikiams sprendimams, susijusiems su pacientų priežiūra, ir gali lemti nevienodą poveikį gyventojų grupėms, išskirtoms pagal lytį, socialinę lytį, rasinę ar etninę kilmę.

Pamatinis šio modelio principas įtvirtintas Chartijos 1 straipsnyje, kuriame nustatyta, kad žmogaus orumas yra neliečiamas. Šis principas grindžia valdymo reikalavimus, užtikrinančius, kad DI priemonės būtų naudojamos taip, jog būtų išsaugotas žmogaus sprendimų priėmimo vaidmuo ir gerbiama paciento autonomija. Praktikoje tai reiškia, kad turi būti numatytos **apsaugos priemonės, padedančios išvengti pernelyg didelio pasitikėjimo automatizuotomis rekomendacijomis, sudarytos galimybės ginčyti**

³⁵ European Union, [Europos Sąjungos pagrindinių teisių chartija](#), 2012/C 326/02, Strasbourg: EU.

pasitelkiant DI priimtus sprendimus ir įdiegtos klinikinės priežiūros priemonės, užkertančios kelią automatizuotui šališkumui – reiškiniui, kai sprendimus priimančias asmuo pernelyg pasitiki algoritmų pateikiamais rezultatais, o tai gali dar labiau sustiprinti nelygybę, jei DI rekomendacijos sistemai yra mažiau palankios tam tikroms gyventojų grupėms, išskirtoms remiantis detalizuota analize.

Panašiai ir Chartijos 20 bei 21 straipsniai, įtvirtinantys lygybės prieš įstatymą ir nediskriminavimo principus, sudaro teisinį pagrindą aiškiems teisingumo užtikrinimo ir šališkumo mažinimo reikalavimams medicininio DI srityje. Todėl ligoninės, diegiančios DI sistemas, turėtų **įdiegti procesus, leidžiančius stebėti jų veikimą įvairiose pacientų grupėse, atsižvelgiant į lytį, socialinę lytį, rasinę ar etninę kilmę**, taip pat į šių požymių derinius. Nustačius DI sistemos veikimo ar jos rezultatų skirtumus tarp atskirų gyventojų grupių, turėtų būti taikomos korekcinės priemonės, pavyzdžiui, koreguojamas modelis, keičiami klinikiniai procesai arba pasirenkami alternatyvūs sprendimų priėmimo būdai, siekiant užtikrinti, kad DI sistemų diegimas neįtvirtintų esamos struktūrinės nelygybės.

Be to, Chartijos 35 straipsnyje įtvirtinta teisė į aukšto lygio žmonių sveikatos apsaugą dar labiau sustiprina principą, kad DI sistemos gali būti diegiamos tik tuo atveju, jei yra įrodyta, jog jos prisideda prie saugios, veiksmingos ir teisingos sveikatos priežiūros. DI technologijos neturėtų sudaryti naujų kliūčių gauti sveikatos priežiūros paslaugas ar didinti jau egzistuojančios nelygybės, su kuria susiduria istoriškai nepakankamai aptarnaujamos ar kitais aspektais nepalankioje padėtyje esančios pacientų grupės. Todėl jų poveikis šioms pacientų grupėms turėtų būti nuolat vertinamas.

Siekiant užtikrinti šių teisių įgyvendinimą praktikoje, šiame politikos dokumente siūlomas reguliavimo modelis grindžiamas visą **DI sistemos gyvavimo ciklą apimančiu valdymo požiūriu, pagal kurį DI priežiūra organizuojama taikant privalomus sprendimų priėmimo etapus** (*angl. gates*). Šie etapai atspindi Dirbtinio intelekto akte ir kituose sveikatos priežiūrą reglamentuojančiuose teisės aktuose įtvirtintą reguliavimo logiką, kartu suteikdami ligoninėms aišką ir audituojamą DI valdymo sistemą. Modelyje daroma prielaida, kad ligoninė paprastai veikia kaip DI sistemos diegėja (pagal Dirbtinio intelekto aktą) ir sveikatos priežiūros paslaugų teikėja pagal nacionalinius sveikatos priežiūros teisės aktus, tačiau jame taip pat numatyta situacija, kai ligoninė tampa DI sistemos teikėja arba gamintoja, pavyzdžiui, kurdama ir naudodama vidaus reikmėms skirtą DI sistemą.

Ligoninės yra pagrindinės DI sistemų diegėjos ir atsakingos už jų įgyvendinimą praktikoje, o pilietinės visuomenės organizacijos veikia kaip nepriklausomos priežiūros, audito ir žmogaus teisių stebėsenos subjektai, nedalyvaujantys kasdieniame klinikinių sprendimų dėl DI sistemų naudojimo priėmime.

Šis reguliavimo modelis įgyvendinamas laikantis proporcingumo principo, pripažįstant, kad **skirtingų valstybių narių sveikatos priežiūros sistemose veikiančios ligoninės ir**

Finansuojama Europos Sąjungos lėšomis. Tačiau išreiškiamas požiūris ar nuomonė yra tik autoriaus (-ių) ir nebūtinai atspindi Europos Sąjungos ar Europos švietimo ir kultūros vykdomosios įstaigos (EACEA) požiūrį ar nuomonę. Nei Europos Sąjunga, nei EACEA negali būti laikoma už juos atsakinga. Projekto kodas: 101215009 – AEQUITAS – CERV-2024-CHAR-LITI

pilietinės visuomenės organizacijos gali turėti nevienodus techninius, organizacinius ir analitinius pajėgumus. Todėl visoms institucijoms taikomi vienodi privalomi minimalūs reikalavimai, užtikrinantys bazinį saugos, skaidrumo ir teisingumo apsaugos lygį. Kartu pažangesnės priemonės, pavyzdžiui, išplėstinė detalizuota gyventojų grupių analizė, išorės auditai ar pažangūs šališkumo stebėsenos įrankiai, gali būti diegiami palaipsniui, atsižvelgiant į konkrečios institucijos pajėgumus. Tokio pat principo laikomasi ir techniniame priede, kuriame aiškiai atskiriami privalomi minimalūs ir rekomenduojami reikalavimai.

1. **Taikymo srities ir klasifikavimo etapas.** Prieš įsigyjant arba kuriant DI sistemą, institucijos turi nustatyti jos teisinį statusą. Tai apima vertinimą, ar sistema priskirtina medicinos priemonių programinei įrangai pagal MPR ar IVPR, ar ji laikoma didelės rizikos DI sistema pagal Dirbtinio intelekto aktą, ar patenka į mažesnės rizikos kategoriją, kuriai vis tiek taikomi DI valdymo reikalavimai. Šiame etape ligoninės taip pat gali nustatyti DI sistemas, kurių numatyta paskirtis gali turėti didesnį poveikį teisingumui arba skirtingoms pacientų grupėms. Tokios sistemos vėlesniuose etapuose turėtų būti išsamiau vertinamos lyčių ir rasinio šališkumo požiūriu. Praktinių pavyzdžių, padedančių išvengti netinkamo DI sistemų teisinio klasifikavimo ir sudarančių prielaidas vėlesniam šališkumo bei teisingumo vertinimui, pateikiama MDCG gairėse dėl programinės įrangos kvalifikavimo ir klasifikavimo.³⁶ Už DI sistemų klasifikavimą ir jų teisinio statuso nustatymą atsako ligoninės, o pilietinės visuomenės organizacijos šiame etape teikia nepriklausomas konsultacijas dėl galimų su teisingumu susijusių rizikų, tačiau sprendimų priėmimo įgaliojimų neturi.
2. **DI sistemų įsigijimo ir sutarčių sudarymo etapas.** Praktikoje didžiausią įtaką DI valdymui ligoninės gali daryti DI sistemų įsigijimo etape. Todėl šiame etape jos turėtų užtikrinti, kad tiekėjai sutartimis įsipareigotų viso DI sistemos gyvavimo ciklo metu valdyti su teisingumu susijusias rizikas, įskaitant lyčių ir rasinį šališkumą. Tiekėjai turėtų pateikti dokumentus, patvirtinančius atitiktį taikomiems reguliavimo reikalavimams, įskaitant, kai taikoma, atitikties vertinimo rezultatus, DI sistemų veikimo validavimo duomenis ir kitą informaciją, reikalingą diegėjams jų pareigoms pagal Dirbtinio intelekto aktą vykdyti. DI sistemų įsigijimo sutartyse taip pat turėtų būti aiškiai apibrėžytas atsakomybės pasidalijimas tarp teikėjo ir diegėjo, nustatant reikalavimus dėl skaidrumo, dokumentacijos, veiklos žurnalų tvarkymo, žmogaus vykdomos priežiūros ir pagalbos po DI sistemos įdiegimo. Pagrindinis šio etapo tikslas – sutartiniuose įsipareigojimuose įtvirtinti šališkumo prevencijos ir atskaitomybės

³⁶ Medical Device Coordination Group, [Programinės įrangos kvalifikavimo ir klasifikavimo pagal Reglamentą \(ES\) 2017/745 dėl medicinos priemonių \(MPR\) ir Reglamentą \(ES\) 2017/746 dėl in vitro diagnostikos medicinos priemonių \(IVPR\) gairės](#), Brussels: MDCG, 2019.

principus, sudarant prielaidas saugiam ir teisingam DI sistemų diegimui. Už DI sistemų įsigijimą ir sutarčių sudarymą atsako ligoninės, o pilietinės visuomenės organizacijos gali peržiūrėti su įsigijimu susijusius dokumentus teisingumo ir žmogaus teisių požiūriu bei teikti neprivalomo pobūdžio rekomendacijas.

3. **Validavimo etapas.** Trečiajame etape atliekamas DI sistemos validavimas, siekiant užtikrinti, kad prieš pradedant ją taikyti klinikiniuose procesuose ji būtų nepriklausomai įvertinta. Validuojant turėtų būti vertinamas ne tik bendras DI sistemos veikimas, bet ir jos rezultatai skirtingose pacientų grupėse, remiantis detalizuota analize, ypatingą dėmesį skiriant lyčiai, socialinei lyčiai, rasinei ir etninei kilmei. Šis etapas ypač svarbus, nes medicininio DI sistemų veikimas skirtingose gyventojų grupėse gali skirtis, jei mokymo duomenys nėra pakankamai reprezentatyvūs arba neatsižvelgiama į skirtumus tarp sistemos kūrimo ir jos taikymo aplinkos.³⁷ Vertinimo metu taip pat turėtų būti dokumentuojamos taikytos rizikos mažinimo priemonės ir aiškiai įvardijamos išliekančios su teisingumu susijusios rizikos, kad ligoninės turėtų pakankamai informacijos priimti pagrįstą sprendimą dėl DI sistemos diegimo. Už validavimo atlikimą arba jo organizavimą ir rezultatų įvertinimą, priimant sprendimą dėl DI sistemos diegimo, atsako ligoninės. Pilietinės visuomenės organizacijos gali atlikti nepriklausomą DI sistemos veikimo skirtingose gyventojų grupėse ir galimų šališkumo rizikų peržiūrą, tačiau nedalyvauja priimant sprendimus dėl validavimo rezultatų.
4. **Diegimo etapas.** Kai DI sistema sėkmingai pereina validavimo etapą, priimamas sprendimas dėl jos diegimo. Šiame etape ligoninės turi užtikrinti, kad būtų sudarytos sąlygos saugiam, teisingam ir atsakingam DI sistemos naudojimui. Tai apima už žmogaus vykdomą priežiūrą atsakingų ir tinkamai parengtų darbuotojų paskyrimą, galimybės gydytojams nepaisyti DI sistemos rekomendacijų arba prireikus sustabdyti jos veikimą užtikrinimą, taip pat naudotojų supažindinimą su DI sistemos numatyta paskirtimi, jos taikymo ribomis ir veikimo ypatumais skirtingose gyventojų grupėse, nustatytais remiantis detalizuota analize. Sveikatos priežiūros specialistai turėtų būti informuoti apie galimus DI sistemos rezultatų skirtumus, susijusius su lytimi, socialine lytimi, rasine ar etnine kilme, o pacientai – apie tai, kad jų sveikatos priežiūros procese naudojamos DI grindžiamos sprendimų priėmimo priemonės. Visa atsakomybė už DI sistemos diegimą, jos integravimą į klinikinius procesus ir diegimo priežiūrą tenka ligoninėms. Pilietinės visuomenės organizacijos diegimo procese nedalyvauja, tačiau gali peržiūrėti pasirengimo diegti DI sistemą dokumentaciją teisingumo ir žmogaus teisių užtikrinimo požiūriu.

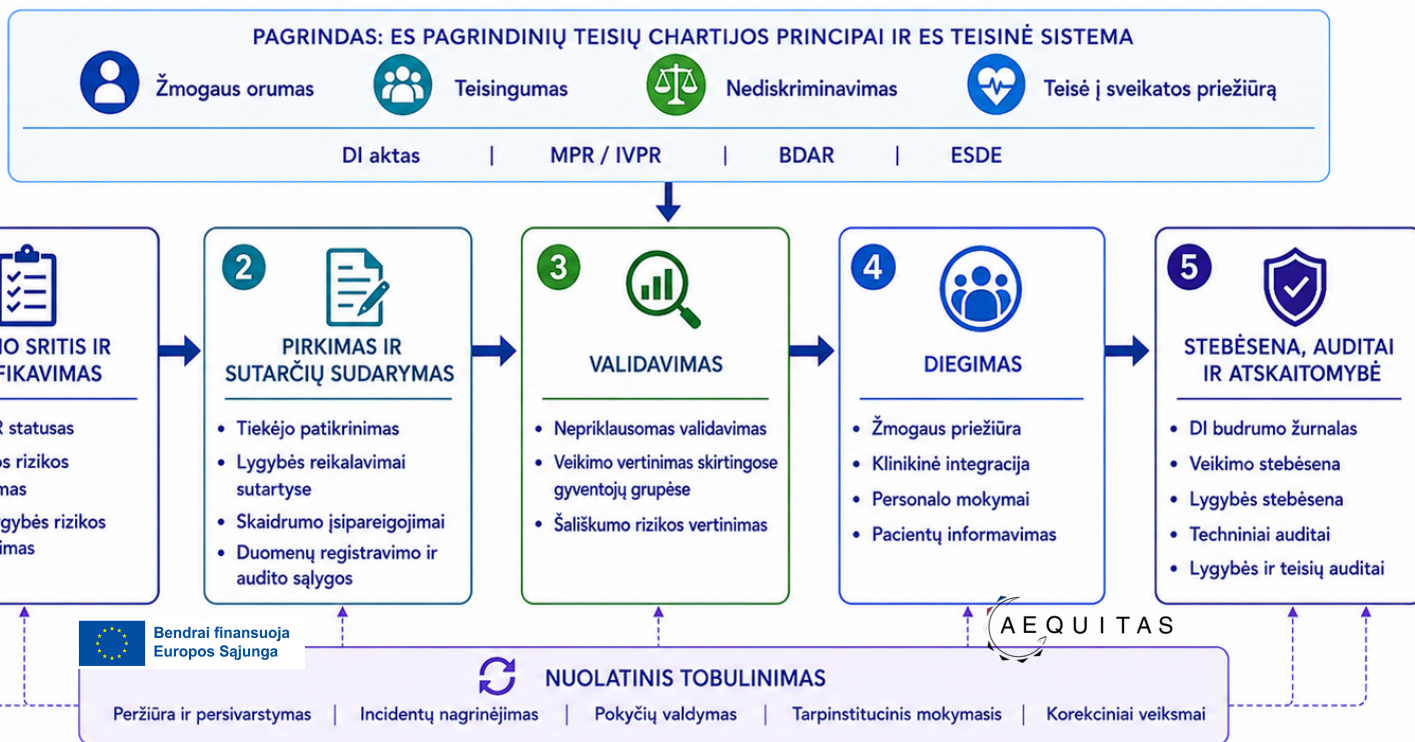
³⁷ Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H., [Teisingumo užtikrinimas taikant mašininį mokymąsi siekiant didinti teisingumą sveikatos priežiūros srityje](#), *Annals of Internal Medicine*, 169(12), 866-872, 2018.

5. **Stebėsenos, audito ir atskaitomybės etapas.** Šis paskutinis etapas apima visą laikotarpį po DI sistemos įdiegimo. DI sistemos turėtų būti nuolat stebimos, vertinant jų bendrą veikimą, rezultatus skirtingose gyventojų grupėse, remiantis detalizuota analize, taip pat registruojant saugos ir su teisingumu susijusius incidentus. Ligoninės turėtų tvarkyti DI sistemų veiklos žurnalus, taikyti standartizuotas incidentų registravimo procedūras, leidžiančias nustatyti galimus lyčių ar rasinio šališkumo atvejus, ir periodiškai iš naujo vertinti DI sistemų veikimą, atliekant pakartotinį validavimą. Nustačius, kad DI sistema lemia nelygybę ar šališkumo riziką, kurios neįmanoma veiksmingai sumažinti, ji turėtų būti koreguojama, jos naudojimas laikinai sustabdomas arba ji turėtų būti visiškai pašalinama iš naudojimo. Šio etapo metu surinkta informacija turėtų būti naudojama ne tik vidaus valdymo sprendimams priimti, bet, kai tikslinga, ir platesniam keitimuisi patirtimi sveikatos priežiūros sistemoje, siekiant užkirsti kelią panašių šališkumo atvejų pasikartojimui. Ligoninės atsako už nuolatinę stebėseną, veiklos žurnalų tvarkymą, incidentų registravimą ir rizikos mažinimo priemonių įgyvendinimą. Pilietinės visuomenės organizacijos atlieka nepriklausomus auditus, vertina sistemos su teisingumu susijusias rizikas skirtingose įstaigose ir padeda atkreipti dėmesį į nuolat pasikartojančius nelygybės atvejus, tačiau nedalyvauja priimant operacinius sprendimus dėl DI sistemų keitimo ar jų naudojimo nutraukimo.

Visi šie DI sistemos gyvavimo ciklo etapai kartu leidžia pagrindinių teisių principus paversti konkrečiais DI valdymo procesais, padedančiais veiksmingai valdyti su lyčių ir rasiniu šališkumu susijusias rizikas klinikinėse DI sistemose.

1 pav. AEQUITAS: ES pagrindinių teisių chartija grindžiamo DI valdymo modelio schema

AEQUITAS ES PAGRINDINIŲ TEISIŲ CHARTIJA GRĮSTA MEDICININIO DI VALDymo SCHEMA



Praktinės politikos priemonės, veiklos įrankiai ir DI valdymo dokumentai ligoninių ir pilietinės visuomenės organizacijų specialistams

Siekiant užtikrinti veiksmingą šio politikos dokumento siūlomo reguliavimo modelio taikymą realioje sveikatos priežiūros praktikoje, numatyta parengti praktinių priemonių rinkinį, kurį būtų galima naudoti įvairiose institucijose ir pritaikyti skirtingiems nacionaliniams kontekstams. Modelyje pabrėžiama standartizuotų procedūrų, dokumentų ir šablonų svarba, siekiant užtikrinti, kad tiek ligoninės, tiek pilietinės visuomenės organizacijos DI sistemas valdytų nuosekliai ir vienodais principais. Toliau pateikiamas priemonių rinkinys parengtas atsižvelgiant į Dirbtinio intelekto akto reikalavimus, MPR ir IVPR taikymo gaires bei kitų sektorių gerąją valdymo praktiką.

DI sistemų įsigijimo ir sutarčių sudarymo etapas

- **DI sistemų įsigijimo ir sutarčių sudarymo dokumentų rinkinys.** ES DI viešųjų pirkimų bendruomenė yra parengusi pavyzdines sutarčių sąlygas (atskiras didelės ir mažesnės rizikos DI sistemų versijas), kurios buvo atnaujintos atsižvelgiant į Dirbtinio intelekto akto reikalavimus. Viešojo sektoriaus institucijos aiškiai skatinamos jomis naudotis, kartu pateikiant ir jų taikymo gaires.³⁸ Siekiant įvertinti ir mažinti su lyčių bei rasiniu šališkumu susijusias rizikas medicininio DI srityje, šias pavyzdines sąlygas galima pritaikyti kaip ligoninėms skirtą atskirą sutarties priedą, kuriame turėtų būti numatyti bent šie reikalavimai:
 - (1) standartizuotas mokymo duomenų sudėties aprašymas, įskaitant duomenis apie pasiskirstymą pagal lytį, socialinę lytį, rasinę ir etninę kilmę, siekiant įvertinti duomenų reprezentatyvumą ir galimą šališkumo riziką;
 - (2) informacijos apie duomenų kilmę atskleidimas, nurodant, ar naudojami duomenų rinkiniai yra nuosavybiniai ar viešai prieinami, kokios yra jų prieinamumo sąlygos ir kokią įtaką tai turi skaidrumui, nepriklausomam validavimui bei lyčių ir rasinio šališkumo vertinimui;
 - (3) informacijos apie DI sistemos veikimą skirtingose gyventojų grupėse, remiantis detalizuota analize, ir validavimo rezultatų atskleidimas, įskaitant pagal lytį, socialinę lytį, rasinę ar etninę kilmę suskirstytus rezultatus, nustatytų veikimo skirtumų dokumentavimą, kuriant DI

³⁸ European Commission, [Atnaujintos ES pavyzdinės sutarties sąlygos dėl dirbtinio intelekto](#), Brussels: European Commission, 2025.

sistemą taikytas šališkumo mažinimo priemonės ir validavimo bandymų rezultatus (įskaitant konkrečiai gydymo įstaigai reikšmingus teisingumo aspektus);

- (4) reikalavimai dėl veiklos žurnalų tvarkymo ir audituojamumo, užtikrinant, kad DI sistema galėtų generuoti naudojimo žurnalus, reikalingus nepriklausomam auditui atlikti ir galimam šališkumui ar nevienodam poveikiui skirtingoms gyventojų grupėms įvertinti;
- (5) reikalavimai dėl DI sistemos atnaujinimų ir pokyčių valdymo, įpareigojant teikėją informuoti ligoninę apie modelio atnaujinimus, pakartotinį modelio mokymą ar klinikinių procesų pakeitimus, taip pat dokumentuoti galimą šių pokyčių poveikį DI sistemos veikimui skirtingose gyventojų grupėse;
- (6) įsipareigojimai dėl reagavimo į su šališkumu susijusius incidentus, įpareigojant teikėją pranešti apie nustatytas klaidas ar nelygybę, darančią poveikį pacientų grupėms, išskirtoms pagal lytį, socialinę lytį, rasinę ar etninę kilmę, ir pateikti numatytų taisomųjų priemonių planą;
- (7) aiškų atsakomybės pasidalijimą, apibrėžiant teikėjo ir diegėjo pareigas, susijusias su Dirbtinio intelekto akto reikalavimų laikymusi, taip pat su teisingumo užtikrinimu ir su šališkumu susijusių rizikų valdymu.

Validavimo ir pasirengimo diegti etapai

- **Sveikatos duomenų valdymo ir reprezentatyvumo politika.** Siekdamas mažinti lyčių ir rasinį šališkumą, visos DI sistemas diegiančios ligoninės ir pilietinės visuomenės organizacijos turėtų patvirtinti oficialią politiką, užtikrinančią, kad DI sistemoms mokytis, papildomai apmokytis ar diegti naudojami duomenys būtų reprezentatyvūs, išsamūs ir tiksliai atspindėtų aptarnaujamą pacientų grupę. Tokia politika turėtų apimti demografinės duomenų sudėties vertinimą, nepakankamai atstovaujamų pacientų grupių nustatymą ir dokumentavimą bei, prireikus, priemonių šiai problemai spręsti numatymą. Ji turėtų būti taikoma tiek pačios įstaigos sukauptiems, tiek iš išorės gaunamiems duomenų rinkiniams, nepriklausomai nuo to, ar jie yra nuosavybiniai, ar viešai prieinami. Be to, politika turėtų numatyti pareigą tikrinti DI sistemų veikimo rezultatus skirtingose gyventojų grupėse, remiantis detalizuota analize, ir vertinti tiekėjų taikomas šališkumo mažinimo priemones. Toks požiūris atitinka Dirbtinio intelekto akto 10 straipsnio reikalavimus dėl šališkumo rizikos nustatymo ir mažinimo bei dera su gerąja praktika, aprašoma mokslinėje literatūroje apie medicininio DI šališkumo mažinimą visais DI sistemos gyvavimo ciklo etapais – nuo duomenų rinkimo iki stebėsenos po sistemos įdiegimo.³⁹ Ši politika nėra savarankiškas valdymo dokumentas, o veikia bendras politikos lygmens

³⁹ Char, D.S., Shah, N.H., & Magnus, D., [Mašininio mokymosi diegimas sveikatos priežiūroje: etinių iššūkių sprendimas](#), *New England Journal of Medicine* 378: 981-983 (2018).

dokumentas, kuriuo grindžiamos kitos šiame modelyje numatytos priemonės. Ji taikoma kaip vienas iš pasirengimo DI sistemos diegimui principų ir sudaro pagrindą vėlesnei nuolatinei stebėsenai, auditui ir atskaitomybei.

- Kibernetinio saugumo baziniai reikalavimai ir DI sistemoms pritaikytas grėsmių modelis.** Ligoninėse naudojamos DI sistemos gali būti pažeidžiamos ne tik įprastų kibernetinių atakų, tokių kaip duomenų užteršimas, priešiškieji pavyzdžiai ar modelio apėjimo atakos, bet ir manipuliacijų ar klaidų, galinčių sustiprinti lyčių ar rasinį šališkumą, pavyzdžiui, neproporcingai paveikti nepakankamai mokymo ar diegimo duomenyse atstovaujamas demografinės grupės. Dirbtinio intelekto akto 15 straipsnyje nustatytas reikalavimas užtikrinti DI sistemų atsparumą neteisėtiems jų rezultatų ar veikimo pakeitimams, todėl būtina taikyti specialiai DI sistemoms skirtas apsaugos priemonės. Dėl šios priežasties ligoninės turėtų nustatyti bazinius kibernetinio saugumo reikalavimus ir parengti DI sistemoms pritaikytą grėsmių modelį, kuriame kartu su techninėmis grėsmėmis būtų aiškiai įvertintos ir su šališkumu susijusios rizikos. Šis nuolat atnaujinamas dokumentas turėtų atlikti dvejopą funkciją – būti ir politikos dokumentu, ir praktiniu įrankiu. Prieš diegiant DI sistemą jis galėtų būti naudojamas kaip kontrolinis sąrašas, padedantis įsitikinti, kad įgyvendintos kibernetinio saugumo priemonės atitinka MPR ir IVPR reikalavimus, organizacijos informacijos saugumo valdymo sistemą (pvz., ISO/IEC 27001) bei į šališkumo riziką orientuotus grėsmių modeliavimo ir testavimo planus.⁴⁰ Po DI sistemos įdiegimo šis dokumentas taip pat padėtų vykdyti nuolatinę stebėseną, leidžiančią laiku nustatyti naujas technines ir su šališkumu susijusias rizikas bei imtis prevencinių veiksmų. Tokios politikos įgyvendinimas padėtų sustiprinti DI sistemų atsparumą, užtikrinti teisingumą DI grindžiamų sprendimų priėmimo procese ir laikytis Dirbtinio intelekto akto (9–15 straipsniai) bei MPR ir IVPR reikalavimų.
- Individualios DI sistemos byla.** Tai standartizuotas dokumentas, rengiamas kiekvienai DI sistemai prieš ją diegiant sveikatos priežiūros įstaigoje. Byloje turėtų būti glaustai aprašyta DI sistemos numatyta paskirtis, jos integravimas į klinikinius procesus, validavimo rezultatai ir veikimo rodikliai, ypatingą dėmesį skiriant DI sistemos veikimui skirtingose pacientų grupėse, remiantis detalizuota analize, pagal lytį, socialinę lytį, rasinę ar etninę kilmę, taip pat nustatytiems veikimo skirtumams ir taikytoms jų mažinimo priemonėms. Joje taip pat turėtų būti aprašytos žmogaus vykdomos priežiūros priemonės, aiškiai nurodyta, kurie sprendimai išlieka išimtinėje sveikatos priežiūros specialistų kompetencijoje, ir pažymėtos visos kontraindikacijos, reikšmingos pacientų saugai ir teisingumui užtikrinti. Pateikdama standartizuotą ir į šališkumo riziką

⁴⁰ Medical Device Coordination Group, [MDCG 2019-16 ed. – Medicinos priemonių kibernetinio saugumo gairės](#), Brussels: European Commission, 2020.

orientuotą DI sistemos apžvalgą, ši byla didina skaidrumą, stiprina vidaus valdymą ir palengvina reguliavimo priežiūrą, užtikrindama, kad priimant sprendimus dėl DI sistemos diegimo būtų aiškiai įvertintos galimos neteisingų rezultatų rizikos.

- **Bendras poveikio pagrindinėms teisėms ir poveikio duomenų apsaugai vertinimo dokumentų rinkinys (FRIA ir DPIA).** Viešosios ligoninės, diegiančios didelės rizikos DI sistemas, vis dažniau privalo vertinti ne tik galimą poveikį pagrindinėms teisėms, bet ir riziką asmens duomenų apsaugai. Šių vertinimų sujungimas į vieną standartizuotą šabloną, atitinkantį Dirbtinio intelekto akto (27 straipsnis), BDAR ir poveikio duomenų apsaugai vertinimo (DPIA) reikalavimus, gali sumažinti administracinę naštą ir kartu užtikrinti, kad būtų nuosekliai įvertinti visi esminiai aspektai: paveiktos gyventojų grupės, galimos žalos, rizikos mažinimo priemonės ir skundų nagrinėjimo mechanizmai. Ypatingas dėmesys turėtų būti skiriamas galimai nelygybei, galinčiai paveikti pacientų grupes, išskirtas pagal lytį, socialinę lytį, rasinę ar etninę kilmę.

Tai periodiškai atnaujinamas audito dokumentas, taikomas pasirengimo DI sistemos diegimui etape kaip teisingumo ir pagrindinių teisių audito priemonė. Jame sistemingai ir dokumentuotai vertinama galima rizika pagrindinėms teisėms ir asmens duomenų apsaugai. Šis dokumentas atlieka ne nuolatinės stebėsenos, o valdymo kontrolės priemonės funkciją. Nors pasikeitus rizikoms ar DI sistemos veikimo sąlygoms jis turėtų būti atnaujinamas, pagrindinė jo paskirtis – prieš DI sistemos diegimą ir jos naudojimo metu užtikrinti sistemingą, dokumentuotą ir peržiūrimą galimo poveikio vertinimą. Tai padeda užtikrinti atitiktį tiek Dirbtinio intelekto akto, tiek BDAR reikalavimams, taip pat laiku nustatyti ir mažinti su lytimi, socialine lytimi, rasine ar etnine kilme susijusio šališkumo riziką.

Siekiant sustiprinti poveikio vertinimo žmogaus teisių aspektą, rengiant šių vertinimų metodiką turėtų būti numatytos prasmingos konsultacijos su pilietinės visuomenės organizacijomis, pasitelkiant jų kompetenciją žmogaus teisių srityje. Pavyzdžiui, Europos nacionalinių žmogaus teisių institucijų tinklas (ENNHRI) ragina Europos Komisiją poveikio vertinimo šablonus papildyti metodinėmis gairėmis ir numatyti viešas konsultacijas, kad poveikio pagrindinėms teisėms vertinimai (FRIA) būtų veiksmingi ir tinkamai įvertintų teisingumo aspektus lyties, socialinės lyties, rasinės ir etninės kilmės požiūriu.⁴¹

⁴¹ European Network of National Human Rights Institutions (ENNHRI), [Pareiškimas dėl veiksmingo poveikio pagrindinėms teisėms vertinimo \(FRIA\) užtikrinimo pagal ES Dirbtinio intelekto aktą](#), Brussels: ENNHRI, 2025.

- **Pokyčių valdymo politika.** Ši politika reglamentuoja visus DI sistemų pakeitimus, įskaitant modelio pakartotinį mokymą, programinės įrangos atnaujinimus ir klinikinių procesų pakeitimus. Organizacijos turėtų aiškiai apibrėžti, kokie pakeitimai laikomi esminiais ir dėl to reikalauja pakartotinio validavimo arba papildomo patvirtinimo, siekiant užtikrinti, kad jie nesukeltų nenumatytų rizikų pacientų priežiūrai. Ypatingas dėmesys turėtų būti skiriamas tam, kaip šie pakeitimai gali paveikti lyčių ir rasinį teisingumą, įskaitant galimus DI sistemos veikimo skirtumus skirtingose gyventojų grupėse, nustatytus remiantis detalizuota analize, ar nevienodus rezultatus pagal demografines charakteristikas. Ligoninės taip pat turėtų užtikrinti, kad gydytojai ir kiti susiję darbuotojai būtų laiku informuojami apie pakeitimus, galinčius turėti įtakos skirtingų pacientų grupių gydymo rezultatams, kad prireikus galėtų taikyti tinkamas priežiūros ir rizikos mažinimo priemonės. Integruodamos teisingumo principus į pokyčių valdymo procesą, ligoninės gali užtikrinti, kad DI grindžiamas sprendimų priėmimas klinikinėje praktikoje išliktų ne tik saugus, bet ir teisingas.
- **Kokybės vadybos sistema (KVS) diegėjams, kurie kartu yra ir DI sistemų teikėjai.** Svarbi šio modelio dalis skirta DI sistemoms, kurias pačios kuria ligoninės ar mokslinių tyrimų institucijos. Viešosios sveikatos priežiūros įstaigos, ypač akademiniai medicinos centrai, vis dažniau pačios kuria arba pritaiko DI priemones, todėl tokiais atvejais jos faktiškai tampa DI sistemos gamintojomis arba teikėjomis. Tai reiškia, kad joms tenka ir papildomos reguliavimo pareigos. Europos lygmens gairėse dėl įstaigų vidaus reikmėms gaminamų medicinos priemonių pabrėžiama, kad sveikatos priežiūros įstaigos turėtų taikyti kokybės vadybos sistemą ir tvarkyti techninę dokumentaciją, įrodančią, jog jų pačių sukurti sprendimai tenkina konkrečius pacientų poreikius, kurių negali patenkinti rinkoje prieinamos komercinės priemonės.⁴² Kokybės vadybos sistema turėtų užtikrinti, kad DI sistemų veikimas būtų vertinamas skirtingose pacientų grupėse, remiantis detalizuota analize, pagal lytį, socialinę lytį, rasinę ar etninę kilmę, kad būtų dokumentuojamos taikomos šališkumo mažinimo priemonės, o žmogaus vykdomos priežiūros mechanizmai aiškiai užtikrintų pacientų saugą ir teisingumą. Net ir tais atvejais, kai vidaus reikmėms sukurtoms medicinos priemonėms taikomos tam tikros reguliavimo išimtys, pagrindinių teisių apsaugos ir bendrieji DI valdymo reikalavimai išlieka visiškai taikytini. Todėl kokybės vadybos sistema tampa viena svarbiausių priemonių, padedančių integruoti teisingumo principus į medicininio DI sistemų kūrimą, validavimą ir diegimą.

⁴² Medical Device Coordination Group, [MDCG 2023-1 – Gairės dėl sveikatos priežiūros įstaigoms taikomos išimties pagal MPR ir IVPR](#), Brussels: European Commission, 2023.

Stebėsenos, audito ir atskaitomybės etapas

- **DI sistemų veikimo skirtingose gyventojų grupėse ir teisingumo vertinimo ataskaita.** Kadangi DI sistemų veikimo skirtumai gali lemti nevienodus sveikatos priežiūros rezultatus, ligoninės turėtų ne tik reikalauti, kad tiekėjai prieš DI sistemų įsigijimą pateiktų atitinkamą informaciją, bet ir pačios reguliariai rengti vidaus analizes, parodančias, kaip DI sistemos veikia skirtingose pacientų grupėse pagal jų demografines charakteristikas. Šiose analizėse pagrindinis dėmesys turėtų būti skiriamas DI sistemų veikimo rezultatams, suskirstytiems pagal lytį, socialinę lytį, rasinę ar etninę kilmę, kai tai leidžiama pagal teisės aktus ir yra praktiškai įmanoma, taip pat turi būti pagrindžiami visi neišvengiami veikimo skirtumai. Nustačius tokius skirtumus, organizacijos turėtų dokumentuoti taikomas rizikos mažinimo priemones ir nuolat vertinti jų veiksmingumą. Toks požiūris atitinka Europos DI reguliavimo sistemoje stiprėjančius lūkesčius, kad organizacijos aktyviai vertintų ir mažintų su šališkumu susijusias rizikas tiek duomenų rinkiniuose, tiek DI sistemų rezultatuose. Dirbtinio intelekto akto 10 straipsnyje aiškiai nustatytas reikalavimas vertinti ir mažinti šališkumo riziką duomenų rinkiniuose, taip pat, taikant griežtas apsaugos priemones, leidžiama ribota specialių kategorijų duomenų tvarkymo apimtis šališkumui nustatyti.

DI sistemų veikimo skirtingose gyventojų grupėse ir teisingumo vertinimo ataskaita pirmiausia yra nuolatinės stebėsenos priemonė. Ji numato nuolatinį duomenų rinkimą, reguliarią DI sistemų veikimo stebėseną skirtingose demografinėse grupėse ir nuolatinį taikomų rizikos mažinimo priemonių atnaujinimą, kai nustatomi nevienodi rezultatai. Tai leidžia sveikatos priežiūros įstaigoms laiku nustatyti su šališkumu susijusias rizikas ir imtis tinkamų veiksmų joms mažinti. Kartu svarbiausios šios ataskaitos išvados, įskaitant nustatytus veikimo skirtumus ir taikytų rizikos mažinimo priemonių veiksmingumą, turėtų būti apibendrinamos ir įtraukiamos į teisingumo ir pagrindinių teisių audito ataskaitą, taip užtikrinant didesnę skaidrumą ir viešąją atskaitomybę.

- **Dviejų lygmenų audito sistema, apimanti techninį auditą ir teisingumo bei pagrindinių teisių auditą.**

Ligoninėse taikoma audito sistema turėtų užtikrinti pusiausvyrą tarp išsamumo ir praktinio įgyvendinamumo, būti suprantama tiek darbuotojams, tiek visuomenei. Šiame modelyje auditas suprantamas kaip formalizuotas, periodiškai atliekamas DI sistemos atitikties reikalavimams, veikimo ir poveikio pagrindinėms teisėms vertinimas. Jis papildo nuolatinę DI sistemos stebėseną, tačiau jos nepakeičia.

Finansuojama Europos Sąjungos lėšomis. Tačiau išreiškiamas požiūris ar nuomonė yra tik autoriaus (-ių) ir nebūtinai atspindi Europos Sąjungos ar Europos švietimo ir kultūros vykdomosios įstaigos (EACEA) požiūrį ar nuomonę. Nei Europos Sąjunga, nei EACEA negali būti laikoma už juos atsakinga. Projekto kodas: 101215009 – AEQUITAS – CERV-2024-CHAR-LITI

Techninis auditas: Šis audito lygmuo apima konfidencialų vidaus vertinimą, kurio metu atliekami išsamūs techniniai patikrinimai, įskaitant kibernetinio saugumo testavimą, DI sistemos atsparumo vertinimą, klinikinio veikimo validavimą ir duomenų valdymo reikalavimų laikymosi vertinimą. Ypatingas dėmesys turėtų būti skiriamas DI sistemos veikimui skirtingose gyventojų grupėse, remiantis detalizuota analize, pagal lytį, socialinę lytį, rasinę ar etninę kilmę, vertinant, ar modelio rezultatai sistemingai nėra mažiau palankūs kuriai nors gyventojų grupei. Šis auditas turėtų būti atliekamas laikantis atitinkamų Dirbtinio intelekto akto (9–15 straipsniai), o prireikus – ir MPR bei IVPR reikalavimų.

Teisių ir teisingumo auditas: Antrasis audito lygmuo skirtas skaidrumui, atskaitomybei ir šališkumo mažinimui užtikrinti. Jo metu turėtų būti dokumentuojami nustatyti gydymo rezultatų netolygumai, testuojant DI sistemą įtrauktos gyventojų grupės, žinomi sistemos apribojimai ir skundų nagrinėjimo mechanizmai, kartu aiškiai įvertinant poveikį lyties, socialinės lyties, rasinės ir etninės kilmės požiūriu. Audite taip pat turėtų būti pateikiami įrodymai, kad žmogaus vykdoma priežiūra yra veiksminga ir leidžia nustatyti bei ištaisyti šališkus DI sistemos rezultatus. Remiantis užimtumo srityje taikomu šališkumo auditų ir viešo informacijos atskleidimo modeliu⁴³ (*bias audit & disclosure*), visuomenei galėtų būti skelbiama vieša audito santrauka, kurioje būtų pateikiamos pagrindinės audito išvados, taikytos rizikos mažinimo priemonės ir įgyvendinti patobulinimai, neatskleidžiant nuosavybinės ar kitos konfidencialios informacijos. Šiam audito lygmeniui taip pat priskiriamas bendras poveikio pagrindinėms teisėms ir poveikio duomenų apsaugai vertinimo dokumentų rinkinys (FRIA ir DPIA), padedantis užtikrinti išsamų pagrindinių teisių, asmens duomenų apsaugos ir teisingumo skirtingų demografinių grupių atžvilgiu vertinimą.

- **DI budrumo žurnalas.** DI sistemų stebėseną po jų įdiegimo taip pat turėtų būti grindžiama DI budrumo žurnalu, parengtu pagal farmakologinio budrumo sistemų, taikomų vaistų reguliavimo srityje, principus.⁴⁴ Sveikatos priežiūros specialistai turėtų turėti galimybę registruoti incidentus, beveik įvykusius incidentus ir netikėtus atvejus, susijusius su DI grindžiamais klinikiniais sprendimais. Žurnale turėtų būti numatyti atskiri laukai informacijai apie galimus skirtumus, susijusius su lytimi, socialine lytimi, rasine ar etnine kilmė,

⁴³ New York City Department of Consumer and Worker Protection, [Automatizuotos sprendimų priėmimo priemonės įdarbinimo srityje \(AEDT\)](#), New York: City of New York, 2025.

⁴⁴ European Medicines Agency, [Geros farmakologinio budrumo praktikos \(GVP\) VI modulis. Įtariamų nepageidaujamų reakcijų rinkimas ir pranešimas apie jas \(2 red.\)](#), London: EMA, 2017.

fiksuoti. Surinkti pranešimai turėtų būti sistemingai analizuojami siekiant nustatyti pasikartojančias tendencijas, įskaitant nevienodą poveikį atskiroms demografinėms grupėms, kad organizacijos galėtų laiku identifikuoti saugos signalus, su teisingumu susijusias problemas ar šališkumo tendencijas. Prireikus, ypač tais atvejais, kai incidentai rodo galimą nevienodą pacientų gydymą ar riziką saugomoms gyventojų grupėms, apie juos turėtų būti nedelsiant pranešama DI sistemos teikėjui, atsakingiems ligoninės valdymo organams ir kompetentingoms priežiūros institucijoms. DI budrumo žurnalo duomenys taip pat turėtų būti naudojami atliekant periodinius teisingumo ir pagrindinių teisių auditus, taip prisidedant prie nuolatinio DI sistemų tobulinimo ir šališkumo mažinimo klinikinėje praktikoje.

- **Bendradarbiavimas tinklo lygmeniu ir keitimasis informacija apie incidentus.**

Europos pilietinės visuomenės organizacijų ir viešųjų ligoninių tinklas, sukurtas pagal AEQUITAS sistemą, veikia kaip bendra DI valdymo platforma, užtikrinanti vienodą terminiją ir standartizuotus pranešimų apie su lyčių ir rasiniu šališkumu DI sistemose susijusius incidentus bei jų vertinimo rezultatus principus. Šis tinklas skatina keitimąsi patirtimi tarp skirtingų įstaigų ir gerosios praktikos sklaidą nustatant, mažinant ir šalinant nelygybę, remiantis didelės rizikos sektoriuose ir farmakologinio budrumo srityje taikomais principais. Kaupiant ir susiejant duomenis apie šališkumo incidentus, DI sistemų veikimą skirtingose gyventojų grupėse, remiantis detalizuota analize, ir taikomas rizikos mažinimo priemonės, tinklas sudaro galimybes ligoninėms ir pilietinės visuomenės organizacijoms nustatyti sisteminės nelygybės tendencijas ir koordinuotai taikyti tikslines intervencijas. Kadangi Europos sveikatos duomenų erdvė (EHDS) reglamentuotomis sąlygomis sudaro galimybes naudoti sveikatos duomenis tiek pirminiais tikslais (sveikatos priežiūros paslaugų teikimui), tiek antriniais tikslais (moksliniams tyrimams ir viešosios politikos formavimui), šis tinklas, tobulėjęs ESDE teisiniam reguliavimui ir priimant jo įgyvendinimo aktus, galėtų savo dalijimosi duomenimis ir tarpinstitucinio mokymosi praktiką derinti prie EHDS reikalavimų. Tai padėtų užtikrinti teisingesnį, skaidresnį ir nuoseklesnį DI sistemų diegimą įvairiose sveikatos priežiūros įstaigose.

- **Mokymų modulis sveikatos priežiūros specialistams ir PVO ekspertams.**

Siekiant mažinti automatizavimo šališkumą ir nelygybę, galinčią atsirasti taikant klinikinį DI, AEQUITAS reguliavimo modelyje numatyti privalomi mokymai, skirti suteikti sveikatos priežiūros specialistams žinių ir įgūdžių, reikalingų suprasti DI sistemų ribotumus, kritiškai vertinti jų pateikiamus rezultatus ir prireikus nepaisyti DI sistemos rekomendacijų arba sustabdyti jos naudojimą, jei kyla rizika, kad ji gali lemti šališkus ar neteisingus sprendimus. Mokymuose ypatingas dėmesys skiriamas su lyčių ir rasiniu šališkumu susijusių

rizikų atpažinimui, pagrindinių teisių principams ir etiniams aspektams, įskaitant teisingumą, nediskriminavimą ir pacientų įtrauktį.

Šis modulis laikytinas pagalbine reguliavimo modelio įgyvendinimo priemone, padedančia užtikrinti, kad už DI sistemų priežiūrą atsakingi darbuotojai turėtų pakankamai kompetencijų tinkamai vykdyti stebėsenos, validavimo ir atskaitomybės procedūras, vadovaudamiesi teisingumo principu. Mokymų programoje taip pat turėtų būti numatytos praktinės atvejų analizės, skirtos atpažinti ir mažinti skirtumus, susijusius su lytimi, socialine lytimi, rasine ar etnine kilme.

Išvados

DI suteikia reikšmingų galimybių gerinti sveikatos priežiūros rezultatus, tačiau nesant tinkamo valdymo jis gali atkartoti ar net sustiprinti nelygybę, susijusią su lytimi, socialine lytimi, rasine ar etnine kilme, taip pat kitas sisteminio šališkumo formas. Šiame politikos dokumente siūlomas reguliavimo modelis pagrindinių teisių principus paverčia konkrečiomis apsaugos priemonėmis, taikomomis visais DI sistemos gyvavimo ciklo etapais – nuo įsigijimo ir validavimo iki diegimo, stebėsenos ir atskaitomybės, ypatingą dėmesį skiriant gydymo rezultatų skirtumų tarp skirtingų pacientų grupių nustatymui ir mažinimui. Įgyvendindamos DI valdymo sistemą, grindžiamą aiškiai apibrėžtais gyvavimo ciklo etapais, politikomis, praktiniais įrankiais ir valdymo dokumentais, ligoninės ir pilietinės visuomenės organizacijos gali sistemingai vertinti, mažinti ir dokumentuoti rizikas, įskaitant nevienodą poveikį marginalizuotoms pacientų grupėms. Reguliavimo modelyje derinamos kitų sektorių gerosios praktikos, Europos Sąjungos teisės aktų reikalavimai ir tarptautiniai DI valdymo principai, taip suteikiant sveikatos priežiūros specialistams priemones atsakingai valdyti DI sistemas ir užtikrinti, kad sveikatos priežiūra būtų teisinga, saugi ir skaidri visiems pacientams, nepriklausomai nuo jų lyties, socialinės lyties, rasinės ar etninės kilmės ar kitų požymių.