Vilniaus universiteto Teisės fakulteto

Privatinės teisės katedra

Gytautės Peseckaitės-Kibickienės

V kurso, civilinės ir verslo teisės

studijų šakos studentės

Magistro darbas

Mašininiu mokymu grįstų sprendimų keliami svarbiausi dabartiniai teisiniai iššūkiai (Most Important Current Legal Challenges of Machine Learning Based Solutions)

Vadovas: dr. S. Drazdauskas

Recenzentė: doc. dr. L. Mikalonienė

Vilnius

2022

ANTOTACIJA IR PAGRINDINIAI ŽODŽIAI

Šiame darbe analizuojami mašininiu mokymusi grįstų sprendimų keliami iššūkiai susiję su asmens privatumo apsauga, vartotojų apsauga ir atsakomybe už žalą. Analizuojamos doktrinoje pateikiamos nuomonės, kaip mašininis mokymasis ir dirbtinis intelektas paveikia susiformavusią reguliacinę aplinką per mašininio mokymosi algoritmų naudojamus duomenis ir jų apdorojimą profiliuojant ar reitinguojant asmenis, ir kiek suteikiama teisinė apsauga gali būti paveikta savarankiškai veikiančių algoritmų.

Pagrindiniai žodžiai: mašininis mokymaisi, dirbtinis intelektas, algoritmai, duomenys, išvestiniai duomenys, profiliavimas, reitingavimas.

This paper analyses the challenges posed by machine-learning-based solutions related to the protection of personal privacy, consumer protection, and liability for damages. Opinions in the doctrine are analysed on how machine learning and artificial intelligence affect the formed regulatory environment through the data used by machine learning algorithms and their processing in profiling or ranking individuals, and the extent to which the legal protection provided can be influenced by self-acting algorithms.

Key words: machine learning, artificial intelligence, algorithms, data, derived, inferred data, profiling, ranking.

Turinys

ĮŽANGA

Temos aktualumas. Mašininis mokymasis (*angl. Machine learning, ML*) kaip dirbtinio intelekto (DI) technologija šiuolaikiniame pasaulyje tampa neatsiejama dalimi įvairiausio spektro sprendimų, nuo paprastų algoritmų, kurie filtruoja elektroninį paštą ir rūšiuoja laiškus pagal jų turinį, iki sudėtingų algoritmų, kurie naudojami savaeigiuose automobiliuose. „Mašininio mokymosi metodai, be kita ko, naudojami paieškos sistemose, kurios automatiškai taiso rašybos klaidas, taip pat sudėtingesnėse srityse, tokiose kaip sukčiavimo prevencija, rizikos analizė, klientų elgesio įžvalgos tobulinimas ir medicinos mokslo tobulinimas“ (Algorithms and human rights, 2017, p. 7). Kiekvienas asmuo šiandien neišvengiamai susiduri su viena ar kita dirbtinio intelekto technologija, kurio viena iš posistemių yra mašininis mokymasis. Dirbtinio intelekto technologijos sparčiai vystosi, o investicijos ir inovacijų skatinimas dirbtinio intelekto technologijų srityje tampa daugumos valstybių strateginiais prioritetais. Europos Sąjungos (ES), Jungtinių Amerikos Valstijų, Jungtinės Karalystės ir kitų šalių praktikoje, kai dirbtinio intelekto technologijomis paremti sprendimai paliečia valstybinį ir privatų gyvenimą, iškilo klausimai ir poreikis pradėti reguliuoti dirbtinio intelekto technologijomis paremtus sprendimus, atsižvelgiant į šiomis technologijomis keliamą riziką ir galimus teisinius iššūkius. Šališkumo, diskriminavimo, atsakomybės klausimai keliami nagrinėjant mašininiu mokymusi paremtus sprendimus asmenų atžvilgiu, nes dirbtinio intelekto technologijos, o ypač mažai kontroliuojamos ir pačios apsimokančios, gali kelti dar neatpažintus teisnius klausimus, kurie kyla iš Europos Sąjungos pagrindinių teisių ir laisvių apsaugos, saugos ir sveikatos, gero administravimo principų. Vis didėjančios duomenų, naudojamų mašininio mokymosi ir dirbtinio intelekto sprendimams, apimtys formuoja naujas ekonomines rinkas ir keičia požiūrį į asmens duomenų vertę, kai atsikaitymas už paslaugos internetinėje erdvėje tarp tiekėjo ir vartotojo vykdomas ne pinigais, o asmens duomenimis.

Europos Sąjungoje pastaruosius penkerius metus stebimas naujo reguliavimo, susieto su skaitmeninėmis rinkomis, poreikis ir iniciatyvos, nes taip vadinama ketvirtoji pramonės revoliucija, paremta technologiniu proveržiu, verčia peržiūrėti pamatinių Europos Sąjungos vertybių apsaugos struktūrą. Per pastaruosius metus atsirado naujos arba buvo peržiūrėtos galiojančios Europos Sąjungos direktyvos ir reglamentai, dėl ko atsiranda nauji reikalavimai skaitmeninėje erdvėje plėtojamiems verslams ir viešosioms paslaugoms. 2021 m. išleistas pasiūlymas Europos Parlamento ir Tarybos Reglamentui,

kuriuo nustatomos suderintos dirbtinio intelekto taisyklės (Dirbtinio intelekto aktas), kuris harmonizuos dirbtinio intelekto technologijoms taikomus reikalavimus, nustatys priežiūros mechanizmą ir atlieps keliamus teisinius ir etinius iššūkius. Jungtinės Amerikos Valstijos, Jungtinė Karalystė ir tarptautinės organizacijos nuosekliai eina prie dirbtinio intelekto technologijomis paremtų technologijų reguliavimo klausimų įgyvendinimo ir svarbiausių teisinių iššūkių sprendimo, kuriuos sukelia šios technologijos.

Darbo tikslas. Šiame darbe siekiama atsakyti į klausimą, kokius svarbiausius teisnius iššūkius kelia mašininio mokymosi technologijomis pagrįsti sprendimai asmens privatumui, vartotojų apsaugai ir civilinės atsakomybės sąlygoms ir kokiomis priemonėmis siūloma tuos iššūkius spręsti. Atsakyti į keliamus klausimus keliami uždaviniai:

1. Išanalizuoti mašininio mokymosi sampratą ir veikimo principą, aptarti pagrindinius mašininio mokymosi technologijos elementus ir apibendrinti jų keliamus iššūkius;
2. Aptarti mašininiu mokymu priimamų sprendimų spektrą ir įvertinti keliamus tokių sprendimų teisnius iššūkius;
3. Aptarti pagrindinius, svarbiausius probleminius mašininiu mokymusi priimamų sprendimų aspektus asmens privatumui ir duomenų apsaugai;
4. Išanalizuoti mašininio mokymosi poveikį vartotojų informuotam pasirinkimui ir aptarti kokiais reguliavimo pokyčiais sprendžiamos šios problemos;
5. Išanalizuoti ir įvertinti moksliniame diskurse vyraujančias pozicijas dėl mašininio mokymosi priimamų sprendimų keliamų civilinės atsakomybės problemų;
6. Suformuoti išvadas ir pateikti pasiūlymus dėl mašininiu mokymusi pagrįstų sprendimų reguliavimo efektyvumo ir poreikių.

Darbo objektas – mašininio mokymosi technologija pagrįsti sprendimai ir kaip šie sprendimai veikia asmens privatumą ir duomenų apsaugos klausimus, vartotojų apsaugą bei kaip mašininio mokymosi sprendimai paveikia civilinės atsakomybės sąlygų nustatymą. Mašininio mokymosi sprendimų taikymo spektras ir probleminių aspektų apimtis gali apimti įvairias teisės šakas ir institutus, nuo konstitucinės doktrinos iki pamatinių žmogaus teisių. Koncentruojantis į naujausius reguliavimo pokyčius šiame darbe pasirenkama apsiriboti tais teisės aspektais, kuriuos paliečia mašininio mokymosi sprendimai asmens privatumui ir asmens duomenų apsaugai, vartotojų apsaugai ir civilinei atsakomybei. Įvairūs viešosios teisės probleminiai aspektai, kuriuos šiandien

paliečia dirbtinio intelekto technologijos nėra šio darbo objektai, siekiant paliesti tuos civilinės teisės aspektus, kurie daro tiesioginę įtaką kiekvienam asmeniui.

Darbe plačiau nagrinėjami duomenų, naudojamų mašininio mokymosi sprendimas probleminiai aspektai, nes mokymosi duomenys ir jų kokybė šiandien yra pagrindinė dedamoji norint taikyti mašininio mokymosi algoritmus. Be to, kad duomenys turi būti teisėtai gaunami, jie turi būti apsaugoti nuo neteisėto jų panaudojimo (kibernetinis saugumas), bet kartu ir patys duomenys turi būti kokybiški bei atsakingai parenkami. Atitinkamai mašininio mokymosi algoritmai šiandien tampa rinkos žaidėjais, nes dėl jų autonominio veikimo atsiranda grėsmė nekontroliuoti sprendimų priėmimo. Vienu aspektu galima įpareigoti taisyklėmis ir sureguliuoti privalomas sąlygas algoritmų darbui, kitu aspektu atsiranda rizika dėl per didelio spaudimo kontrolei prarasti inovacijų vystymosi greitį, todėl ypač svarbu surasti balansą tarp teisėto ir neteisėto algoritmų veikimo. Pirmoji probleminė sritis yra susieta su mašininio mokymosi sprendimams naudojamų duomenų kokybe ir teisėtumu bei įžvalgomis, kokius asmens privatumo apsaugos iššūkius kelia mašininio mokymosi priimam sprendimai. Šioje dalyje nagrinėjamas klausimas, ar asmens duomenų apsaugos reguliavimas šiandien apsaugo asmens privatumą, ypač kai kalbame apie profiliavimą ir išvestinių duomenų panaudojimą mašininio mokymosi algoritmų veikime. Antroji problema skirta vartotojų teisių apsaugos aspektams, kai mašininio mokymosi algoritmai pritaikomi individualizuotai kainodarai ar agresyviai rinkodarai. Nors Europos Sąjungos mastu vykstanti reguliavimo transformacija stipriai keičia ne tik pačios vartotojų apsaugos, bet ir civilinėje apyvartoje dalyvaujančių teisių sampratą, mašininio mokymosi algoritmų priimami sprendimai turėdami „juodosios dėžės“ (*angl. black box*) efektą, gali riboti efektyvų tokios apsaugos taikymą. Paskutinioji darbo dalis skiriama išanalizuoti, kokias kliūtis teisių gynybai ir civilinės atsakomybės taikymui gali kelti mašininiu mokymusi paremti sprendimai. Dėstomojoje darbo dalyje yra paliečiami intelektinės nuosavybės, komercinių paslapčių, sutarčių teisės aspektai, tačiau jie atskirai kaip objektai giliau darbe nėra nagrinėjami.

Tyrimo metodas. Darbas yra analitinio pobūdžio, apimantis doktrinos ir teisės aktų analizę. Dokumentų analizės metodas darbe naudojamas kai nagrinėjami teisės aktai ir teisės aktų projektai, moksliniai tyrimai ir doktrina. Pasiremiama atliktais Europos Parlamento, Europos Komisijos inicijuotais tyrimais ir analitine medžiaga. Įvairūs autoriai pateikia įžvalgas ir vertinimus apie aktualias problemines mašininio mokymosi sprendimų sritis, kurias sistemiškai bandoma integruoti problematikai išryškinti. Lingvistinis metodas taikomas aiškinant teisės aktų normas ir jų reikšmę, pateikiant ir

analizuojant naudojamas sąvokas ir jų taikymą. Sisteminis metodas naudojamas nagrinėjant teisės aktų normų įgyvendinimą bei santykį su probleminiais klausimais. Teleologinis metodas naudojamas aiškinat, kokie yra reglamentavimo tikslai ir siekiai, ypač detaliau analizuojant teisės normas reglamentuojančias asmens duomenų saugą ir vartotojų apsaugą.

Darbo originalumas. Mašininio mokymosi technologijos Lietuvoje yra nagrinėjamos daugiausia iš algoritmų panaudojimo galimybių analizės skirtingose prekių, paslaugų, tyrimų srityse, atskiri teisiniai aspektai yra paliečiami per dirbtinio intelekto sampratą ir panaudojimo pasekmes priimamiems sprendimams, pavyzdžiui savaeigiai automobiliai, autonominiai prietaisai, tačiau teisiniai mašininiu mokymusi pagrįstų sprendimų iššūkiai moksliniuose darbuose nagrinėjami labai mažai. Didelį įdirbį per tyrimus šioje srityje yra padariusi Europos Sąjungos Aukšto lygio ekspertų grupė dirbtinio intelekto klausimais, kuri parengė *„Patikimumo DI etikos gaires“*, atsiranda moksliniai darbai ir disertacijos dirbtinio intelekto etikos ir teisės problematikos temomis, tačiau dėl spartaus dirbtinio intelekto technologijų vystymosi ir visgi riboto realaus šios technologijos galimybių panaudojimo probleminiai aspektai iškyla nuolat.

Lietuvoje yra skelbiama magistro darbų apie dirbtinį intelektą ir su juo susietus teisnius aspektus, ypač viešosios teisės srityje, tačiau nuoseklios mokslinės ir teismų praktikos šioje srityje dar nėra.

Šaltiniai. Pagrindiniai šaltiniai šiame darbe yra norminiai teisės aktai, Lietuvos Respublikos civilinis kodeksas, Europos Sąjungos reglamentai ir direktyvos, Europos Sąjungos teisės aktų projektai ir jų lydimoji medžiaga, kurioje moksliniais tyrimais yra pagrindžiami svarbiausi keliami iššūkiai dirbtinio intelekto srityje ir kaip juos siūloma spręsti reguliavimu. Aiškinant mašininio mokymosi sampratą ir teisės problemas pagrindiniai šaltiniai yra 2020 m. vasario 19 d. Europos Komisijos *Baltoji knyga. Dirbtinis intelektas. Europos požiūris į kompetenciją ir pasitikėjimą*. COM(2020) 65, taip pat Europos tarybos 2018 metais išleista studija, *Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications* kurią atliko interneto tarpininkų ekspertų grupė (MSI-NET). Algoritmų keliamoms problemos identifikuoti svarbiausias šaltinis Europos Parlamento 2020 metų studija *Artificial intelligence: How does it work, why does it matter, and what we can do about it?.*

Duomenų apsaugos ir klausimais svarbiausi naudoti šaltiniai yra Europos parlamento studija: *The impact of General Data Protection Regulations (GDPR) on*

artificial intelignce 2020 m., taip pat Sandra Wachter ir Brent Mittelstadt darbas *„A Right to Reasonable Inferences: re-thinking data protection lae in the age of big data and AI“.*

Nagrinėjant vartotojų teisių apsaugos problematiką pasirinkti šaltiniai yra Rory Macmillan 2019 m. publikacija *„Big data, machine learning, consumer protection and privacy“*, taip pat Mateusz Grochowski, Agnieszka Jabłonowska, Francesca Lagioia & Giovanni Sartor 2021 m. publikacija „*Algorithmic Transparency and Explainability for EU Consumer Protection: Unwrapping the Regulatory Premises“.*

Civilinės atsakomybės problematikai nagrinėti svarbiausi šaltiniai yra Europos Komisijos 2019 metų Atsakomybės ir naujųjų technologijų ekspertų grupės parengtas darbas „*Liability for artificial intelligence and other emerging digital technologies“,* Maria L. Montagnani ir Mirta Cavallo 2021 m. publikacija *Notre Dame Journal of International & Comparative Law „"Liability and Emerging Digital Technologies: An EU Perspective"* ir Yaniv Benhamou and Justine Ferland 2020 m publikacija „A*rtificial intelligence & damages: assessing liability and calculating the damages“* knygai „*Leading Legal Disruption: Artificial Intelligence and a Toolkit for Lawyers and the Law“.*

Taip pat darbe remiamasi atliktais moksliniais tyrimais, pavyzdžiui Anastasia Siapka disertacija *„The Ethical and Legal Challanges of Artificial Intelligence: the EU response to biased and discriminatory AI“* ir straipsniais iš mokslinių straipsnių duomenų bazių, skelbiamomis oficialiomis ataskaitomis ir pranešimais.

I. MAŠININIO MOKYMOSI TECHNOLOGIJA, VEIKIMO PRINCIPAI IR IŠŠŪKIAI TURINTYS ĮTAKOS TEISĖS KLAUSIMAMS

Šioje darbo dalyje analizuojama mašininio mokymosi technologija, jos samprata ir veikimo principai. Taip pat aptariami svarbiausi mašininio mokymosi technologijos elementai, kurie sukelia esminius klausimus ir riziką. Suprantant patį mašininio mokymosi principą galima išskirti ir pagrindinius klausimus, kurie kyla vertinant mašininiu mokymusi priimtus sprendimus ir jų poveikį tiek dirbtinio intelekto, mašininio mokymosi sprendimų algoritmų kūrėjų, tiek ir jų paveiktų asmenų teises ir pareigas.

1.1.Mašininio mokymosi samprata ir veikimo principai

Mašininis mokymasis kaip technologija turi keletą naudojamų apibrėžimų, tačiau dažniausiai apibrėžiama, kad „mašininis mokymasis yra viena iš DI posistemių, kai algoritmai mokomi nustatyti tam tikrus dėsningumus iš duomenų rinkinio, pagal kuriuos nustatoma, kokių veiksmų reikia konkrečiam tikslui pasiekti“ (Baltoji knyga, 2020, p. 17). Kalbant paprasčiau, tai yra matematinis modelis, kuriam yra nustatomas tikslas atrasti tendencijas duomenyse ir pateikti rezultatą.

Mašininis mokymasis nėra naujas terminas. Ar mašinos gali galvoti, dar 1950 metais klausė britų matematikas Alan‘as Turing‘as savo darbe „Skaičiavimo technika ir intelektas“ (*angl. Computing machinery and intellingence),* kur analizavo, ar mašina, kompiuteris, galėtų veikti taip, kad įtikintų žmogų, jog ji gali mąstyti (Turing, 1950). Savo darbe Turing‘as aiškino kaip mašiną galima išmokyti ir kokius lūkesčius jai galima formuoti, vadinant jas mokymosi mašinomis. Išsakyta mintis, kad „svarbi mokymosi mašinos ypatybė yra ta, kad jos mokytojas labai dažnai nežino, kas vyksta viduje, nors jis vis tiek gali tam tikru mastu numatyti savo mokinio elgesį” (Turing, 1950), yra šiandien diskutuojama praktiškai, kai mašininio mokymosi sprendimai yra neatsiejami nuo bet kokios modernios technologijos ar skaitmeninės paslaugos mūsų kasdienėje veikloje. Savo esamą ideologiją mašininis mokymasis įgavo dutūkstantaisiais metais, kai kompiuterių mokslininkai rėmėsi išskirtiniu žmogaus intelekto bruožu – gebėjimu mokytis, kuris buvo suvokiamas kaip gebėjimas panaudoti patirtį elgsenai gerinti (Siapka, 2018, p. 19). „Automatizavimas yra viena iš pagrindinių savybių, susijusių su algoritminiu sprendimų priėmimu. Automatizuotų skaičiavimo sistemų gebėjimas pakeisti žmones vis daugiau atvejų yra pagrindinė praktinio algoritmų įgyvendinimo savybė“ (Algorithms and human rights, 2017, p. 5).

Duomenų mokslininkai sukūrė mašininio mokymosi metodiką, kuri leidžia algoritmams mokytis iš savęs paties ir pagerinti esamą veikimą. „Žmonių pakeitimo automatizuotomis skaičiavimo sistemomis priežastys dažniausiai gali būti siejamos su didelio masto duomenų apdorojimo, sprendimų priėmimo greičio, apimties ir masto problemomis ir daugeliu atvejų lūkesčiais, kad klaidų lygis bus mažesnis, palyginti su žmonėmis“ (Algorithms and human rights, 2017, p. 6). Tai pasiekiama naudojant mašininio mokymosi metodus, naudojant įvesties duomenis, kad būtų galima nustatyti ir sukurti savo modelius būsimiems rezultatams numatyti.

Mašininis mokymas kaip sąvoka dažnu atveju yra naudojamas dirbtinio intelekto sąvokos kontekste, kaip sinonimai, nes bendroji dirbtinio intelekto sąvoka apima ir mašininį mokymąsi kaip vieną iš dirbtinio intelekto posistemių, kuri yra daugiausiai naudojama. Ilgą laiką mokslininkai diskutavo ir interpretavo kaip būtų galima apibrėžti dirbtinį intelektą ir 2021 metais Europos Sąjungos Dirbtinio intelekto akto pasiūlyme yra apibrėžta dirbtinio intelekto sistemos sąvoka, kaip „programinė įranga, sukurta taikant vieną ar daugiau metodų ir principų ir gebanti pagal tam tikrus žmogaus nustatytus tikslus generuoti išvedinius, pavyzdžiui, turinį, predikcijas (prognozes), rekomendacijas, arba sprendimus, turinčius įtakos aplinkai, su kuria ji sąveikauja“ (2021 m. balandžio 21 dienos Europos Komisijos pasiūlymas, 3 straipsnio 1 punktas). Taigi, mašininis mokymasis yra tik viena iš technologijų, kurios priklauso dirbtinio intelekto technologijų grupei, kaip ir gylusis mokymasis (*angl. Deep learning*, DL) ar natūraliosios kalbos apdorojimas (*angl. Natural language processing*, NLP).

Pagrindinis mašininio mokymosi technologijos elementas yra mokymosi duomenys, kurių pagalba algoritmas atlieka skaičiavimus ir teikia išvadas. „Be duomenų DI neegzistuotų. Daugelio DI sistemų veikimas ir jas naudojant atliekami veiksmai bei priimami sprendimai labai priklauso nuo duomenų rinkinio, naudojamo joms mokyti (Baltoji knyga, 2020, p. 19). Mašininio mokymosi modeliai naudoja didelius duomenis (*angl. Big Data*), kad išmoktų ir pagerintų nuspėjamumą ir našumą automatiškai, naudojant patirtį ir duomenis be programavimo, kad tą darytų žmogus (OECD, 2021, p. 3). Visa tai įgyvendinama per algoritmus. „Priimdami įvairius mokymosi stilius, algoritmai modeliuoja problemas, pagrįstas duomenų rinkiniais, ir sukuria naujus sprendimus, kurių žmogui gali būti neįmanoma suvokti. Iš esmės taikant nuolatinius bandymų ir klaidų metodus, algoritmai aptinka esamų duomenų modelius, nustato panašius būsimų duomenų modelius ir daro duomenimis pagrįstas prognozes“ (Algorithms and human rights, 2017, p. 7). „Priklausomai nuo srities, kuriai jie sukurti, jų klasifikavimo tikslumo ir skaičiavimo išteklių prieinamumo, mašininio mokymosi

modeliai yra sprendimų medžių, logistinės regresijos, neuroninių tinklų arba jų derinių „modelių ansambliuose“ (Burrell, 2016). Mašininis mokymasis turi savo mokymosi procesą, nes “mašina savarankiškai modifikuoja sąveikas savo tinkle taip, kad kiekvieną kartą gaudama įvestį nuosekliai pateiktų laukiamą išvestį” (Reillon, 2018).

Autoriai paprastai išskiria 2 arba 3 mašininio mokymosi tipus, kurie priklauso nuo to, kaip ir iš ko mokymasis mokosi, nes kaip minėta anksčiau, pats technologijos veikimo principas yra vykdyti užduotis pagal sukurtą algoritmą. Mašininis mokymasis pagal veikimo formą skirstomas į:

- Prižiūrimą mokymąsi (*angl. supervised learning*);
- Neprižiūrimą mokymąsi (*angl. unsupervised learning*);
- Sustiprintąjį mokymąsi (angl. *reinforcement learning)* (So, 2020, p. 2).

Prižiūrimo mašininio mokymosi atveju kūrėjas pažymi įvesties duomenis ir jų atitinkamus išėjimus (Siapka, 2018, p. 20), tai reiškia, kad mokomieji duomenys turi informaciją, kurie prižiūri mokymosi procesą (So, 2020, p. 3). Konkretus pavyzdys, mokymas atpažinti vaizdą ir atskirti katiną nuo šuns pagal tam tikrus atributus. Kadangi informacija yra teikiama žmogaus, tai duomenys yra paženklinami, pažymimi žmogaus pagal jo supratimą (apmokantis žmogus priskiria atributą kiekvienam vaizdui ir pasako, kad viename paveiksle yra šuo, kitame katė). „Tai reiškia, kad ženklas (atributas) duotas konkrečiam duomenų pavyzdžiui nebūtinai bus teisingas, nes gali būti paties žmogaus ar duomenų pavyzdžio dviprasmiškumo klaidų“ (So, 2020, p. 3). Prižiūrimo mašininio veikimo atveju žmogus nustato tikslus ir sąlyginai gali kontroliuoti rezultatą.

Tuo tarpu neprižiūrimo mašininio mokymosi atveju įvesties duomenys yra nežymėti ir mašina pati randa tarp jų asociacijas, panašumus (Siapka, 2018, p. 20). Toks veikimas paremtas tikėjimu, kad sugeneruoti duomenys nėra atsitiktiniai ir juose yra užkoduota informacija apie kitus procesus (So, 2020, p. 9). Šiuo atveju neprižiūrimo mokymosi užduotys neturi jokių konkrečių išorinių gairių apie tai, kad jo surasti ryšiai, asociacijos, panašumai yra teisingi ar klaidingi.

Trečiasis veikimo principas, sustiprintasis mokymasis reiškia scenarijų, kaip agentas (čia dažnai kalbama apie tokias mašininio mokymosi sistemas, kurios sąveikauja su aplinka) mokosi bendraudamas ilgą laiką su aplinka, kad pasiektų tam tikrą tikslą. „Sąveika apima agentą, kuris imasi veiksmų, kad pakeistų aplinkos būklę, ir gauna grįžtamąjį ryšį atlygio ir nuobaudų pavidalu iš aplinkos, o jo tikslas paprastai yra maksimaliai padidinti bendrą kaupiamąjį atlygį (efektyvumą)“ (So, 2020, p. 14). Nuo neprižiūrimo mokymosi jis skiriasi tuo, kad juo siekiama maksimaliai padidinti tam tikrą atlygio funkciją, o ne atskleisti paslėptą struktūrą.

Apibendrinant, „kai nėra žmogaus nurodymo, ypač neprižiūrimo mašininio mokymosi atveju, neįmanoma nustatyti, kurias įvesties duomenų ypatybes naudojo mašininio mokymosi modelis, kad pasiektų galutinę išvestį ir kokiais būdais. Todėl prognozės ar paaiškinimai, kaip mašininio mokymosi modeliai apdoroja arba apdoros įvestį, tampa nepasiekiami. Kita vertus, neprižiūrimas mašininis mokymasis yra artimesnis žmogaus intelekto modeliui, o tai padidina jo ilgalaikio pritaikymo galimybes“ (Siapka, 2018, p. 20). Dėl šios priežasties atsiranda „juodosios dėžės“ problematika, kuri kelia daugybę teisinių ir etinių iššūkiu, kuriuos nagrinėja įvairių sričių mokslininkai ir politikos formuotojai. Mašininio mokymosi sprendimams vis labiau plintant praktiniame pritaikyme, yra keliami įvairūs klausimai susiję su viešosios teisės klausimai, nes „dėl daugeliui DI technologijų būdingų ypatybių, tokių kaip neskaidrumas („juodosios dėžės“ reiškinys), sudėtingumas, nenuspėjamumas ir dalinis autonomiškumas, gali būti sunku patikrinti jų atitiktį galiojančių ES teisės aktų, kuriais siekiama apsaugoti pagrindines teises, taisyklėms ir tai gali trukdyti veiksmingai užtikrinti jų vykdymą“ (Baltoji knyga, 2020, p. 12). Iš kitos pusės mašininio mokymosi technologija įsivyrauja privačiuose asmenų santykiuose (kredito rinka, finansų paslaugos, draudimo rinka, medicinos diagnostika ir pan.), kur kyla mašininio mokymosi technologijos priimamų sprendimų rizikos.

Apibūdinant mašininio mokymosi technologijos keliamus iššūkius ir rizikas reikia suprasti, kad „DI yra technologijos, kurias taikant derinami duomenys, algoritmai ir kompiuterijos pajėgumai“ (Baltoji knyga, 2020, p. 2). „Didieji duomenys, mašininis mokymasis ir dirbtinis intelektas suteikia pelningas komercines galimybes ir socialinę naudą per profiliavimą ir automatizuotus sprendimus“ (Macmillan, 2019, p. 17). „Profiliavimas – tai automatizuotas asmens duomenų tvarkymas, siekiant įvertinti, analizuoti ar numatyti galimus asmens interesų, asmeninių pageidavimų, elgesio, darbo rezultatų, ekonominės padėties, sveikatos, patikimumo, buvimo vietos ar judėjimo aspektus (Macmillan, 2019, p. 17). „Automatizuoti sprendimai yra sprendimai, kuriuos priima kompiuterinės apdorojimo sistemos be jokio žmogaus dalyvavimo (neskaitant kodavimo), paprastai remiantis išvadomis, padarytomis profiliuojant naudojant mašininio mokymosi modelius, taikomus dideliems duomenims“ (Macmillan, 2019, p. 17). Nagrinėjant mašininio mokymosi technologiją ir priimamų sprendimų problematiką patys svarbiausi elementai, turintys įtakos teisės klausimams yra **duomenys** ir **algoritmai.**

1.2.Duomenų, naudojamų mašininiam mokymui, problematika

Mašininio mokymosi sprendimų priėmimas visiškai priklauso nuo duomenų, kuriuos naudoja algoritmas. ES Dirbtinio intelekto akto pasiūlymo konstatuojamosios dalies 44 punkte yra pabrėžiama, kad „kokybiški duomenys turi esminę reikšmę užtikrinant daugybės DI sistemų efektyvų veikimą, ypač tais atvejais, kai naudojami su mokymo modeliais susiję metodai“ (2021 m. balandžio 21 dienos Europos Komisijos pasiūlymas). Kadangi duomenys yra pagrindinė „žaliava“ mašininio mokymo sprendimams, tai „mokymų, validavimo ir bandymo duomenų rinkiniai turėtų būti pakankamai aktualūs, reprezentatyvūs ir be klaidų bei išsamūs, kad būtų užtikrinta numatytoji sistemos paskirtis. Jie taip pat turėtų turėti tinkamus statistinius ypatumus, įskaitant duomenis apie asmenis arba asmenų grupes, kurių atžvilgiu numatyta naudoti DI sistemą“ (2021 m. balandžio 21 dienos Europos Komisijos pasiūlymas). Dirbtinio intelekto akto pasiūlymo 3 straipsnyje yra apibrėžiamos duomenų rūšys, kurios naudojamos DI tikslams:

- mokymo duomenys – duomenys, naudojami DI sistemai mokyti pritaikant jos mokymosi parametrus, įskaitant neuroninio tinklo parametrus;
- validavimo duomenys – duomenys, naudojami išmokytai DI sistemai įvertinti ir jos kitiems nei mokymosi parametrams bei mokymosi procesui suderinti, be kita ko, siekiant išvengti persimokymo. Validavimo duomenų rinkinys gali būti atskiras duomenų rinkinys arba pastovi ar kintama mokymo duomenų rinkinio dalis;
- bandymo duomenys – duomenys, naudojami nepriklausomam išmokytos ir validuotos DI sistemos vertinimui atlikti siekiant patvirtinti numatomą tos sistemos veikimą prieš ją pateikiant rinkai arba pradedant naudoti;
- įvesties duomenys – DI sistemai teikiami arba jos tiesiogiai gaunami duomenys, kuriais remdamasi ji generuoja išvedinį.

Tokia duomenų klasifikacija yra itin svarbi, nes būtent duomenų rinkinio problematika kyla, kai nagrinėjame mašininio mokymosi sprendimų priėmimą. Pavyzdžiui, „vertinant algoritmų poveikį žmogaus teisėms, reikia atsižvelgti į tai, kad algoritminių sistemų kūrėjai turi skirtingą diskrecijos lygį nuspręsdami, pavyzdžiui, kokius mokymo duomenis naudoti arba kaip reaguoti į klaidingus teigiamus rezultatus, ir kad algoritmo operatoriaus galia pasireiškia per duomenų rinkinio supratimą, o ne per algoritmo veikimo įžvalgas“ (Algorithms and human rights, 2017, p. 6).

Pati duomenų rinkinio pradžia yra susieta su **duomenų**, naudojamų mašininiam mokymui, **gavimu**. „Duomenų rinkimą palengvina elektroninių duomenų prieinamumas

kaip šalutinis produktas naudojant bet kokią IRT sistemą“ (Sartor, *et al*, 2020, p. 16). „Duomenys dabar yra renkami naudojant daugybę programų ir jutiklių, kurie registruoja vartotojų ryšius, operacijas ir judesius“ (Macmillan, 2019, p. 14). Dalis duomenų yra gaunami iš asmenų per registracijos anketų, dokumentų pildymą siekiant gauti paslaugas (pavyzdžiui bankas, draudimas, socialinės paslaugos ir pan.) ir vykdant operacijas (pavyzdžiui elektroniniai pirkimai, registracijos, pavedimai ir pan.). „Pastaraisiais metais šie duomenų srautai buvo integruoti į pasaulinę tarpusavyje sujungtą duomenų apdorojimo infrastruktūrą, orientuotą į internetą, bet tuo neapsiribojant. Ši infrastruktūra yra universali priemonė bendrauti, pasiekti duomenis ir teikti bet kokias privačias ir viešąsias paslaugas“ (Sartor, *et al,* 2020, p. 16). Šiandien „vienas iš svarbių duomenų šaltinių yra asmens veikla elektroninių ryšių srityje, pvz., socialinė žiniasklaida, el. paštas, naršymas internete ir kita elektroninė veikla, pavyzdžiui, elektroninių operacijų istorija. Didelis duomenų kiekis, gautas naudojant žiniatinklio naršykles ir mobiliųjų telefonų programas, yra renkamas ir bendrinamas netaikant standartinės pasirinkimo politikos“ (Macmillan, 2019, p. 14). Pavyzdžiui atlikta Oxford‘o Universiteto studija parodė, kad ištyrus 1 milijoną Android programėlių rasta, kad apie 90 procentų šių programėlių iš išmaniųjų telefonų siunčia informaciją į *Google*. Taip pat su išmaniaisiais įrenginiais yra atiduodama daug duomenų dėl vietos, kontaktų, vaizdų ir pan., kai vartotojas suteikia leidimą programėlėms naudoti įrenginio duomenis. Taigi duomenų industrija leidžia sekti įvairiausius duomenis, kuriuos seka išmanieji įrenginiai ir juos galima susieti su asmeniu (Macmillan, 2019, p. 14).

Duomenų, naudojamų mašininiam mokymui, apimtys gali būti didinamos, o „duomenys gali būti sujungti didelių duomenų operacijų tikslais. Jie gali būti renkami iš mažmeninės prekybos parduotuvių, kuriose asmuo apsiperka, iš kredito kortelių bendrovių, naudojamų sandoriams, iš Bluetooth aptikimo įrenginių parduotuvėse surinktų duomenų, vaizdo kamerose surinktų asmens vaizdų, vaizdo kamerose surinktų asmens vaizdų, vaizdo kameromis surinktų automobilių numerių, informacijos apie vaistus, surinktos iš vaistų pirkimų, žaislų su sumontuotais mikrofonais ir kameromis įrašų ir daugybės kitų šaltinių“ (Macmillan, 2019, p. 16).

Didžiuosius duomenis apibūdina 3 V koncepcija (*angl. huge Volume, high Velocity and great Variety*) – didelis kiekis, didelis greitis ir didelė įvairovė. Taip pat, kabant apie duomenų panaudojimą dirbtinio intelekto, mašininio mokymosi technologijomis, atsiranda dar 2 dedamosios, t.y, mažas tikslumas *(angl. low Veracity)* ir aukšta vertė *(angl. high Value).* Tai yra būdinga didžiųjų duomenų technologijoms, ir atsižvelgiant į

šias kategorijas algoritmų kūrėjai vadovaujasi kurdami ir analizuodami mašininio mokymosi sprendinius (Macmillan, 2019, p. 14).

Analizuojant duomenų klausimą, mokslininkai daugiausia dėmesio kreipia į duomenų panaudojimą mašininio mokymosi sprendimuose kai yra liečiami privatūs asmenys, t.y. į **sąžiningumą** ir **diskriminavimą** priimat sprendimus, o tai visų pirma kyla iš galimybės panaudoti šališkus duomenis. „Dirbtinio intelekto sistemų naudojami duomenų rinkiniai (kuriais grindžiamas jų mokymasis ar veikimas) per neapdairumą gali atspindėti istoriškai susiformavusias šališkas tendencijas, būti neišsamūs arba netinkamai valdomi. Tokių šališkų tendencijų išlaikymas gali lemti (ne)tiesioginę diskriminaciją. Žalos gali padaryti ir sąmoningas (vartotojų) polinkių išnaudojimas arba nesąžininga konkurencija“ (Pasitikėjimo į žmogų orientuotu ..., 2019). Macmillan sako, kad iš didžiųjų duomenų ir profiliavimo padarytos „išvados ir prognozės pagerina įmonių gebėjimą diskriminuoti vartotojus, siūlydamos jiems produktus ir paslaugas, atitinkančias jų pageidavimus ar poreikius, ir tokiomis kainomis, kurias jos nori mokėti“ (Macmillan, 2019, p. 17) ir tokie pavyzdžiai apima sprendimus, ar suteikti asmeniui kreditą, ar pasiūlyti asmeniui darbą.

Duomenų, naudojamų mašininio mokymosi sprendimuose, šališkumo problema gali atsirasti nuo pačių pirmųjų žingsnių, kai pradedamas kurti dirbtinio intelekto sprendimas. „Mašininio mokymosi algoritmai sukuria modelį iš mokymo duomenų, t. y. istorinių pavyzdžių, kad galėtų numatyti ar priimti sprendimus (Macmillan, 2019, p. 13). „Kai mašininio mokymosi algoritmai mokomi remiantis įvesties duomenimis, kurie yra pagrįsti istoriniais pavyzdžiais, tam tikroms istoriškai nepalankioje padėtyje esančioms gyventojų grupėms jie gali pakenkti. Todėl jie gali atspindėti praeities diskriminaciją, neatsižvelgiant į praeityje kilusias priežastis (pvz., dėl išankstinio nusistatymo ar numanomo šališkumo). Jei tokie ankstesni sprendimai buvo šališki, mašininio mokymosi procesų mokymo duomenys gali išlaikyti arba sustiprinti tolesnį šališkumą“ (Macmillan, 2019, p. 37). „Tendencingumas ir diskriminacija yra bet kuriai visuomeninei ar ekonominei veiklai būdinga rizika. Žmogus, priimdamas sprendimus, nėra apsaugotas nuo klaidų ir neobjektyvumo. Tačiau DI neobjektyvumo padariniai galėtų būti kur kas didesni, dėl jo gali nukentėti ir būti diskriminuojami daug žmonių“ (Baltoji knyga, 2020, p. 12).

Duomenų naudojamų mašininiam mokymuisi klausimas atsiliepia vartotojų apsaugai skaitmeninėse rinkose ir paslaugose, žmogaus privatumo ir asmens duomenų apsaugos sferose. Šie klausimai nagrinėjami toliau.

1.3.Algoritmų keliami iššūkiai

Pagrindinis mašininio mokymosi technologijos variklis yra matematiniai algoritmai, kurie gali būti naudojami žmogaus sprendimams parengti arba nedelsiant juos priimti automatinėmis priemonėmis. Terminas „algoritmas“ yra plačiai taikomas ir turi įvairių reikšmių, priklausomai nuo to, ar jis vartojamas tarp matematikų ir informacijos technologų, komunikacijos ir kultūrinės žiniasklaidos studijose, ar viešumoje, įskaitant politinį ir socialinį diskursą (Algorithms and human rights, 2017, p. 5). Paprastai kalbant, algoritmas yra taisyklių rinkinys, kaip pasiekiamas rezultatas, o mašininio mokymosi atveju turime ir itin nesudėtingus algoritmus (pavyzdžiui, elektroninio pašto filtravimui naudojami algoritmai), ir didelio sudėtingumo, žmogui sunkiai suprantamus algoritmus, kurie ir kelia nemažai klausimų dėl jų patikimumo. „Kad dirbtinis intelektas būtų patikimas, jo algoritmai turi būti pakankamai saugūs, patikimi ir patvarūs, kad visais dirbtinio intelekto sistemos gyvavimo ciklo etapais susidorotų su klaidomis ar neatitikimais ir kad tinkamai reaguotų į klaidingus rezultatus“ (Pasitikėjimo į žmogų orientuotu ..., 2019).

Teisiniame viešajame diskurse yra keliami bendrieji politiniai algoritmų **skaidrumo**, **atskaitingumo** ir **etikos** klausimai, kurių reguliavimą inicijuoja skirtingos jurisdikcijos, atsižvelgiant į algoritmų paplitimą įvairiuose sektoriuose, ypač tuose, kur susiduriama su tiesioginiu (kartais ir netiesioginiu) poveikiu asmeniui per jo turtinių ir neturtinių teisių pažeidimą.

Mašininio mokymosi algoritmai kelia skirtingus skaidrumo klausimus, tačiau Europos Parlamento tyrime (Boucher, 2020, p. 19) yra iškirti 4 pagrindiniai aspektai susiję su algoritmų skaidrumu.

Pirmasis aspektas – mašininio mokymosi algoritmo **paaiškinamumo trūkumas**. Tai yra susiję su tuo, kaip žmogus (ar reguliuotojas) gali suprasti algoritmo sprendimą? Kadangi algoritmas mokosi iš duomenų apie žmogaus patirtį, tai turėtų būti galima žmogui ir paaiškinti kaip ir koks yra priimtas sprendimas. „Tačiau lygiavertis ML algoritmas atlieka milijonus skaičiavimų, vadovaudamasis savo vidine logika. Net jei sprendimai yra geros kokybės, labai sunku – dažnai neįmanoma – paaiškinti sprendimą ar jo logiką taip, kad tai būtų prasminga ekspertams, jau nekalbant apie naudotojus, politikos formuotojus, teisėjus“ (Boucher, 2020, p. 19) ir kitus. „Tai reiškia, kad net jei ML algoritmai priima geros kokybės sprendimus, jų logika gali būti neskaidri. Taip pat „algoritmų priimti sprendimai gali būti pagrįsti neišsamiais, taigi nepatikimais duomenims, jie gali būti neteisėtai paveikti per kibernetinius išpuolius, būti neobjektyvūs

ar tiesiog klaidingi“ (Pasitikėjimo į žmogų orientuotu ..., 2019). Dėl šios priežasties ML algoritmai kartais apibūdinami kaip „juodosios dėžės“ (Boucher, 2020, p. 19).

Antrasis iššūkis yra **priėjimo prie informacijos disbalansas**. Tai reiškia, kad tam tikri asmenys, algoritmo savininkas, kūrėjas, gali pasinaudoti prieigos prie informacijos galimybėmis ir siekti savo komercinių ir strateginių interesų, pavyzdžiui galima pasinaudoti duomenimis apie vartotojus ir numatyti asmenų „norą mokėti“ už prekes. „Kainos gali būti nustatytos viršutinėje diapazono dalyje ir kiekvienam pirkėjui siunčiamos individualios nuolaidos, kurios iš tikrųjų sumažina kainą iki numatomo noro mokėti. Taip sukuriami individualūs kainodaros režimai, pagal kuriuos pirkėjai gali matyti tik lentynos kainą ir jiems pasiūlytą nuolaidą. Jie neturi prieigos prie išsamios informacijos apie tai, kaip buvo apskaičiuota jų kaina, kaip ji lyginama su kitomis arba kokios kainos yra prieinamos kitiems pirkėjams“ (Boucher, 2020, p. 19). Apribodami pirkėjo vaizdą algoritmo savininkai išlaiko strateginį ar komercinį pranašumą.

Dar vienas iššūkis, trečiasis, yra susietas tuo, kad asmenys, ne visuomet **suvokia, kad jie bendrauja su dirbtiniu intelektu**. „Tai gali apimti kliento sąveiką, pvz., tikroviškas pokalbių roboto sąsajas *(angl. chatbot)*, arba užkulisius, susijusius su paraiškų dėl paskolų ar darbo apdorojimu“ (Boucher, 2020, p. 20). Net jeigu asmuo ir žino, kad bendrauja su dirbtinio intelekto agentu, tačiau jam gali būti neleidžiama suprasti, kodėl buvo priimti vienokie ar kitokie sprendimai (dėl sudėtingumo ar informacijos apribojimo). Taip pat naudotojai gali nežinoti, kad jie yra taip sekami su savo asmenine informacija, kad būtų po to ja pasinaudota komerciniais ar ideologiniais tikslais.

Paskutinis iššūkis yra susijęs su poveikiu ir apibrėžtumu, koks yra **numatytas dirbtinio intelekto kūrimo rezultatas.** Ši problematika yra labiau politinė ir socialinė, nes pasitikėjimas dirbtiniu intelektu, mašininio mokymosi algoritmais turi poveikį ilgalaikei perspektyvai. Skirtingi abstrakcijos ir sudėtingumo lygiai sukelia skirtingus neskaidrumo ir skaidrumo iššūkius (O'Neil, 2016, p. 131), tačiau dirbtinio intelekto kūrėjai dažnai pabrėžia teigiamą mašininio mokymo poveikį ir vengia prieštaringų rezultatų, tokių kaip stebėjimo ar karinių programų kūrimas.

Siekiant didinti skaidrumą, mokslininkai ir politikos formuotojai sako, kad „turėtų būti užtikrintas dirbtinio intelekto sistemų **atsekamumas**. Svarbu registruoti ir dokumentuoti tiek sistemų priimtus sprendimus, tiek visą tų sprendimų priėmimo procesą (be kita ko, aprašyti duomenų rinkimo ir žymėjimo procesą ir naudojamą algoritmą). Algoritminį sprendimų priėmimo procesą taip pat turėtų būti galima, kiek įmanoma, **paaiškinti** atitinkamiems asmenims suprantama kalba (Pasitikėjimo į žmogų orientuotu ..., 2019). Problema kyla iš to, kad mašininio mokymo atveju atskleisti ir paaiškinti

procesą pagal algoritmą gali būti sunku arba praktiškai neįmanoma (Macmillan, 2019, p. 50), o netgi ir kai mašininio mokymo sprendimus galima paaiškinti, „tai sprendimų subjektai gali nesutikti su rezultatu“ (Rodrigues, 2020, p. 2).

Algoritmų **atskaitomybės** problematika yra formuluojama per atsakomybės už žalą principą, „pagal kurį asmuo, kuris yra teisiškai ar politiškai atsakingas už žalą, turi pateikti tam tikrą pateisinimą arba kompensaciją. Tačiau asmuo gali būti atsakingas tik tada, kai turi tam tikrą kontrolės laipsnį ta prasme, kad prisidėjo prie žalos arba ją padarė arba gali užkirsti kelią žalai arba ją sumažinti. Teisiškai atskaitomybė pasireiškia per pareigos suteikti teisių gynimo priemonę (pvz., žalos atlyginimą) sąvoką. Įstatymai paprastai nustato atsakomybę asmeniui, kuris gali užkirsti kelią žalai arba sumažinti riziką (pavyzdžiui, per draudimą). Atskaitomybės paskirstymą už algoritminių sprendimų priėmimą apsunkina tai, kad dažnai neaišku, kas turi reikiamą kontrolės laipsnį, kad būtų galima priskirti teisinę ar politinę atskaitomybę“ (Algorithms and human rights, 2017, p. 39).

Atsakomybės klausimas ir problematika bus nagrinėjama toliau, tačiau siekiant algoritmų naudojimo atskaitingumą įteisinti reguliavimu, Europos Komisijos komunikate pasisakyta, kad „turėtų būti nustatyti mechanizmai, kuriais būtų užtikrinta atsakomybė už dirbtinio intelekto sistemas bei jų rezultatus ir su jais susijusi atskaitomybė tiek iki, tiek po jų diegimo. Šiuo atžvilgiu labai svarbi **galimybė atlikti** dirbtinio intelekto sistemų **auditą**, nes vidaus ir išorės auditorių atliekamas dirbtinio intelekto sistemų vertinimas ir tokio vertinimo ataskaitų pateikimas labai padeda užtikrinti technologijos patikimumą. Visų pirma turėtų būti užtikrinta galimybė atlikti pagrindinėms teisėms poveikį darančių prietaikų (*angl. applications*), įskaitant saugumo požiūriu itin svarbias prietaikas, išorės auditą“ (Pasitikėjimo į žmogų orientuotu ..., 2019). Pagal numatytą rizikos mažinimo priemonę „turėtų būti nustatomas, vertinamas, dokumentuojamas ir kuo labiau mažinamas **galimas neigiamas** dirbtinio intelekto sistemų **poveikis**“ (Pasitikėjimo į žmogų orientuotu..., 2019), tačiau praktiškai algoritmų kūrėjas gali susidurti su mašininio mokymosi technologijos specifika, t.y. prisitaikymu prie gaunamų duomenų, kai „prisitaikymas parodomas savarankiško mokymosi algoritmuose, kurie naudoja duomenis naujiems modeliams ir žinioms kurti ir naujoms sprendimų priėmimo taisyklėms generuoti naudojant mašininio mokymosi metodus“ (Williamson, 2016 cituota Algorithms and human rights, 2017, p. 6) ir aplinkybė, kad paaiškinti algoritmo veikimo kartais negali net pats kūrėjas gali apsunkinti tokio reikalavimo įgyvendinimą.

Algoritmų naudojimas kelia ir svarbių etinių klausimų, kurie ne visada yra tiesioginio teisinio reguliavimo sferoje. Ypač tai pasireiškia diskusijose, kaip algoritmas

turėtų nuspręsti hipotetinėje situacijoje, kai, pavyzdžiui, kyla atsakomybės klausimas dėl savaeigio automobilio keliamos grėsmės mažam vaikui, pagyvenusiems žmonėms, daugelio žmonių gyvybėms. „Be tiesioginių reguliavimo mechanizmų, turinčių įtakos algoritmų kodui, taip pat galima apsvarstyti netiesioginius mechanizmus, turinčius įtakos algoritmų kodams. Jie skirti gamybos procesui arba algoritmų gamintojams ir stengiamasi užtikrinti, kad jie žinotų teisinius iššūkius, etines dilemas ir žmogaus teisių problemas, kylančias dėl automatizuoto duomenų apdorojimo ir sprendimų priėmimo metodų“ (Algorithms and human rights, 2017, p. 40). Dirbtinis intelektas gali būti laikomas patikimu tik jei yra plėtojamas ir naudojamas laikantis plačiai pripažįstamų etinių vertybių (Pasitikėjimo į žmogų orientuotu ..., 2019).

Algoritmai, kaip matematiniai modeliai, yra sukuriami žmogaus, kūrėjo, tačiau patys apsimokydami iš savo veikimo tampa autonomiški ir sunkiai valdomi. Autonominis algoritmų veikimas sukelia įvairių teisinių klausimų, ypač susijusių su civiline atsakomybe, tačiau iššūkių kyla ir kitose srityse, kur mašininio mokymosi algoritmai tampa neatsiejama verslo dalimi.

1.4.Mašininio mokymosi sprendimų keliamos teisės problemos

Dirbtinis intelektas, mašininis mokymasis ir kitos skaitmeninės šiuolaikinės technologijos vystosi labai sparčiai, tačiau jos sukuria ne tik naudą visai ekonomikai, tačiau kelia nemažai teisinių ir etinių klausimų, kurie atitinkamai nagrinėjami skirtingose teisės šakose. Pastaruosius penkerius metus ypač daug institucijų, mokslininkių nagrinėja klausimus, kuriuos kelia dirbtinio intelekto, mašininio mokymosi, didžiųjų duomenų panaudojimas.

Valstybės tiek globaliu, tiek nacionaliniu lygiu skatina technologijų vystymąsi, tačiau, reikia pažymėti, kad dirbtinio intelekto, mašininio mokymosi sprendimams įsiliejant į mūsų praktinį gyvenimą, teisiniai klausimai tampa ypač aktualūs. Viešosios teisės sityje nuolatos pabrėžiama, kad „algoritmų ir kitų automatizuotų duomenų apdorojimo metodų naudojimas gali turėti teigiamą ir neigiamą poveikį žmogaus teisių įgyvendinimui ir naudojimuisi jomis“ (Algorithms and human rights, 2017, p. 44). Europos Sąjungos teisėkūros iniciatyvos dirbtinio intelekto sityje pažymi riziką, kad „DI naudojimas gali daryti poveikį pamatinėms ES vertybėms ir pažeisti pagrindines teises, įskaitant teisę į saviraiškos laisvę, susirinkimų laisvę, žmogaus orumą, nediskriminavimą dėl lyties, rasinės arba etninės kilmės, religijos ar tikėjimo, negalios, amžiaus arba seksualinės orientacijos, kai tai aktualu tam tikrose srityse, teisę į asmens duomenų ir

privatumo apsaugą arba teisę į veiksmingą teisminę gynybą ir teisingą bylos nagrinėjimą, taip pat neatitikti vartotojų apsaugos taisyklių. Ši rizika gali kilti dėl to, kad yra bendrųjų DI sistemų projektavimo trūkumų (be kita ko, susijusių su žmogaus atliekama priežiūra), arba dėl to, kad naudojami neobjektyvūs duomenys (pvz., sistemai mokyti naudojami tik arba daugiausia vyrų duomenys, todėl moterų atžvilgiu gaunami neoptimalūs rezultatai) (Baltoji knyga, 2020, p. 11).

Duomenų šališkumo, algoritminio sprendimo skaidrumo, kibernetinio pažeidžiamumo, diskriminacijos, ginčijamumo trūkumo, atsakomybės, privatumo ir daugelį kitų aspektų paliečia mašininio mokymosi sprendimai ne tik viešosios, bet ir privačios, civilinės, teisės srityje. Įvairūs mokslininkai, tyrėjai ir praktikai susiduria su mašinų priimamais sprendimais finansinių paslaugų, kredito rizikos vertinimo, draudimo rizikos vertinimo, medicininių paslaugų, daiktų interneto srityse, kuriose priimti sprendimai gali sukelti materialinių nuostolių, turtinės ir neturtinės žalos, turi įtakos žmogaus privatumui, taip pat vartotojų apsaugai.

Didelis duomenų, naudojamų mašininio mokymosi sprendimams, kiekis šiandien kelia ne tik teisėto jų surinkimo klausimus, tačiau ir teisėto jų panaudojimo darant įvairiausio pobūdžio prognozes, įžvalgas ir sprendimus. Kibernetinių atakų rizika ir duomenų neteisėto užvaldymo problematiką kelia klausimus dėl duomenų neteisėto panaudojimo įvairiais tikslais, kai sistemos gali būti modifikuotos ir sukelti žalos (Motegnani, *et al*, 2021, p. 216).

Mašininio mokymosi algoritmai laikomi intelektinės nuosavybės objektu, todėl diskurse yra keliami klausimai, ar mašininio mokymosi algoritmai, kurie linkę pasikeisti savaime per nuolatinį mokymąsi gali būti apsaugoti autorių teisėmis. Taip pat, kiek intelektinė nuosavybės ir komercinės paslapčių įstatymai gali būti palankūs kūrėjams vis stiprinant apsauginę DI reguliavimo funkciją.

Pažymima, kad besivystant technologijoms vis daugiau teisinių klausimų kyla dėl dirbtinio intelekto teisinio subjektiškumo, t.y. kaip reikėtų vertinti dirbtinį intelektą, kaip sprendimų priėmėją? Bendrinės fizinio ir juridinio asmens teisinės kategorijos kartais įvardijamos kaip nepakankamos pagrįsti dirbtinio intelekto, mašininio mokymosi sprendimų atsakomybę.

Mašininio mokymosi sprendimų klausimus Rodrigues apibendrino ir pateikė juos suvestinėje lentelėje, kur išvardijo ir susiejo 10 teisinių problemų, kurios aktualios dirbtinio intelekto, mašininio mokymosi sprendimų kontekste ir įvardijo sritis, kuriose yra pažeidžiami subjektai (Rodrigues, 2020, p. 8).

1 lentelė. **Problemų priskyrimas pažeidžiamumui**

Teisinė problema	**Labiausiai pažeidžiamų grupių pavyzdžiai**	**Veiksniai, lemiantys / palengvinantys pažeidžiamumą (pavyzdžiai)**
Algoritminio skaidrumo trūkumas	Žmonės, kuriuos atsisakyta įdarbinti, atsisakyta išduoti paskolą, atsisakyta atvykti (įleisti į šalį) / deportuoti, įkalinti, įtraukti į neskraidymo sąrašus arba atsisakyta skirti pašalpas.	Prastas/blogas/nesąžiningas dizainas, netinkami modeliai. Neefektyvus reguliavimas.
Kibernetinio saugumo spragos	MVĮ ir (arba) asmenys, vis labiau pasikliaujantys ir priklausomi nuo technologijų, palaikančių dirbtinį intelektą. Žmonės dirbtinio intelekto valdomose duomenimis grindžiamose ir intensyvios ekonomikos šalyse. Vaikai ir jaunimas	Prastai suprojektuota ir apsaugota technika. Išteklių trūkumas. Investicijos ir priklausomybė nuo DI ir duomenimis pagrįstų technologijų.
Nesąžiningumas, šališkumas ir diskriminacija	Etninės/rasinės/lyties stereotipinės/profiliuotos grupės ir mažumos. Neturtingi / mažas pajamas gaunantys asmenys. Mokiniai turėję žemus pažymius ir neturintys galimybių mokytis	Kūrėjo šališkumas. Nepakankamas etikos klausimas / dėmesys etiškam planavimui / rezultatų testavimo ir patvirtinimo trūkumas. Žmogaus įsikišimo priemonių trūkumas.
Ginčijamumo trūkumas	Duomenų subjektai, kuriems trūksta informacijos, kurios jiems reikia norint pasinaudoti teisėmis.	Trūksta informacijos, reikalingos pasinaudoti teisėmis.
Juridinis asmuo, subjektiškumas, moralinis veiksnys	Žmonės, kurių teisės ir laisvės yra paveiktos / gali konfliktuoti arba konkuruoti.	Neapgalvota politika ir asmens priskyrimas.
Intelektinės nuosavybės problemos	DI kūrinių išradėjai, kūrėjai.	Trūksta nuostatų aiškumo.
Neigiamas poveikis darbuotojams	Jauni darbuotojai Laisvai samdomi/savarankiškai dirbantys darbuotojai.	Perkvalifikavimo ir perkvalifikavimo trūkumas. Nepritaikoma/nelanksti švietimo sistema.
Privatumo ir duomenų apsaugos problemos	Vaikai, neįgalieji ir/ar vyresni asmenys.	Priklausomybė nuo DI ir duomenimis pagrįstų technologijų.
Atsakomybės klausimai, susiję su padaryta žala	Dirbtinio intelekto sistemų naudotojai / asmenys, kuriems daroma žala, pvz., sveikatos / medicinos - neįgalieji, chroniškai sergantys.	Pernelyg didelė priklausomybė nuo dirbtinio intelekto technologijų.
Atsakomybės už žalą trūkumas	DI sistemų naudotojai / asmenys, kuriems daroma žala, ypač civiliai, nukentėję per tarptautines dirbtinio intelekto atakas.	Neatskaitingumo kultūra – lūkesčių stoka ir tokių standartų naudojimas. Išimčių / išlygų naudojimas siekiant išvengti atskaitomybės skatinimo priemonių (viršijančių ir (arba) atitinkančių įstatymus). Jokių ilgalaikių pasekmių.

Šaltinis: Rodrigues Rowena. (2020) *Legal and human rights issues of AI: Gaps, challenges and vulnerabilities.* Prieiga per internetą: www.elsevier.com/locate/jrt

Atsižvelgiant į Rodrigues atliktą apžvalgą, toliau darbe plačiau nagrinėjamos privatumo ir asmens duomenų apsaugos, vartotojų apsaugos ir atsakomybės sritys bei tose srityse aktualūs mašininio mokymosi sprendimų klausimai.

II. MAŠININIO MOKYMOSI SPRENDIMŲ IŠŠŪKIAI ŽMOGAUS PRIVATUMUI IR DUOMENŲ APSAUGAI

Kai kalbama apie dirbtinio intelekto, mašininio mokymosi sprendimų įtaką žmogaus teisėms, vienas esminių elementų yra asmens privatumas, nes dauguma mašininio mokymosi technologijų remiasi duomenimis, kurie yra klasifikuojami kaip asmens duomenys. Privatumas apima informaciją apie asmens privatų gyvenimą, kuri gali būti viešinama tik su asmens sutikimu (išskyrus išimtis, numatytas įstatymu) ir turi būti apsaugota nuo neteisėtų trečiųjų asmenų veiksmų prieš asmenį. „Privatumas apima ir tai, ką asmenys, kurie legaliai surinko duomenis apie mane, gali su tais duomenimis daryti“ (Skiauterienė, 2020, p. 10).

Europos Sąjunga ir kitos šalys (JAV, Jungtinė Karalystė, Kinija) nuosekliai reguliuoja asmens duomenų apsaugos klausimus, sudarydamas prielaidas efektyviam asmens teisių į privatumą apsaugos mechanizmui. Nors Europos Parlamento ir Tarybos reglamentas (ES) 2016/679 (BDAR) įsigaliojo nuo 2018 metų, tačiau reikia pastebėti, kad dirbtinio intelekto, mašininio mokymosi klausimai tiesiogiai reguliavime nebuvo adresuoti, nes daugiau buvo kreipiamas dėmesys į interneto technologijas, todėl jame yra apibrėžtos sąvokos ir naudojami terminai būdingi dabartinio mašininio mokymosi sprendimų problematikai, tokie kaip internetas, socialiniai tinklai, interneto puslapiai ir pan. Visgi tiesiogiai normų dėl dirbtinio intelekto, mašininio mokymosi sprendimų BDAR apibrėžta nėra ir tenka kelti BDAR interpretavimo, aiškinimo aspektus, kurie taikomi mašininio mokymosi sprendimams.

2.1.Asmens duomenų apibrėžties problematika mašininio mokymosi sprendimų atžvilgiu

Nors „ilgiausios ir ilgalaikės diskusijos apie žmogaus teises dėl automatizuoto duomenų tvarkymo ir algoritmų yra susijusios su teise į privatumą (Algorithms and human rights, 2017, p. 12), reikia atkreipti dėmesį, kad „privatumas apima platų sąvokų spektrą“ (Macmillan, 2019, p. 20). Daugelis vertina privatumą kaip saugomą vertybę teisių gynimo prasme ir Macmillan sako, kad kai kurie mokslininkai išreiškia susirūpinimą dėl „individualumo, autonomijos, vientisumo ir orumo“ kaip platesnės asmeninio ir šeimos gyvenimo laisvės idėjų dalies. „Nors privatumas gali reikšti asmens laisvę, kad kiti asmenys netrukdytų asmeniniams pasirinkimams, ypač susijusiems su jo kūnu, didelė privatumo dalis yra susijusi su tuo, ką kas žino apie asmenį, taigi ir su asmens duomenų tvarkymu“ (Macmillan, 2019, p. 20). Iš čia kyla pagrindinė privatumo

pažeidimo rizika, nes privatumas skaitmeniniame kontekste yra susijęs su asmens duomenų surinkimo, naudojimo ir platinimo kontrole. „Duomenų apsaugos įstatymas yra skirtas apsaugoti žmonių privatumą, tapatybę, reputaciją ir savarankiškumą, tačiau šiuo metu jis neapsaugo duomenų subjektų nuo naujos išvadinės analizės rizikos“ (Wachter, Mittelstadt, 2019, p. 5).

Anksčiau minėta, kad dabartinis gaunamų duomenų srautas, naudojamas algoritmų priimamų sprendimų panaudojimui apima įvairiausio pobūdžio duomenis apie asmenį. „Algoritmai naudojami stebint ir profiliuojant asmenis, kurių naršymo modelius įrašo „slapukai“ ir panašios technologijos, tokios kaip skaitmeninis pirštų atspaudų ėmimas, sujungiamas su paieškos užklausomis (paieškos varikliai, virtualūs asistentai). Be to, elgsenos duomenys apdorojami iš išmaniųjų įrenginių, pavyzdžiui, vietos ir kitų jutiklių duomenys, naudojant mobiliųjų įrenginių programėles, todėl kyla vis didesnių iššūkių privatumui ir duomenų apsaugai“ (Tene, Polonetsky, 2012, cituota Algorithms and human rights, 2017, p. 13). Dėl didelio duomenų srauto atsiranda klausimas, kas dabar yra asmens duomenys.

BDAR apibrėžia sąvoką asmens duomenys – bet kokia informacija apie fizinį asmenį, kurio tapatybė nustatyta arba kurio tapatybę galima nustatyti (duomenų subjektas). Fizinis asmuo, kurio tapatybę galima nustatyti, yra asmuo, kurio tapatybę tiesiogiai arba netiesiogiai galima nustatyti, visų pirma pagal identifikatorių, kaip antai vardą ir pavardę, asmens identifikavimo numerį, buvimo vietos duomenis ir interneto identifikatorių arba pagal vieną ar kelis to fizinio asmens fizinės, fiziologinės, genetinės, psichinės, ekonominės, kultūrinės ar socialinės tapatybės požymius (BDAR, 4 str. 1 dalis). Apibrėžimas apima visą spektrą informacijos apie asmenį, kuriuos galima suklasifikuoti į kelias kategorijas:

- Asmens patekti duomenys (pvz. vardas, pašto kodas);
- Duomenys pastebėti apie asmenį (pvz. būvimo vietos duomenys);
- Duomenys gauti iš pateiktų ar pastebėtų duomenų (pvz. rezidavimo šalis gauta pagal pašto kodą);
- Duomenys gauti ir suskaičiuoti iš anksčiau išvardintų duomenų (pvz. kredito reitingas) (Macmillan, 2019, p. 21).

Įvairūs duomenų šaltiniai ir jų sujungimas mašininiam mokymui visų pirma kelia riziką asmens privatumui. „Vartotojai susiduria su privatumo rizika, kai prie jų asmens duomenų gali patekti neleistini vartotojai, jais gali būti piktnaudžiaujama arba jie gali būti naudojami profiliavimui, dėl kurio gali būti daromos subjektyvios išvados apie vartotoją, kurias gali būti sunku patikrinti, ir dėl to gali būti priimti automatizuoti sprendimai,

turintys įtakos asmens gyvenimui“ (Macmillan, 2019, p. 21). Daugeliu atveju rizika yra susieta su tuo, kad duomenys, kurie naudojami mašininio mokymosi sprendimuose nėra nuasmeninti arba yra nuasmeninti taip, kad vis tiek galima identifikuoti asmenį. Teisinė problema kyla dėl to, kad „asmens duomenų, kurie gali būti generuojami ir bendrinami, apimtis dėl didelių duomenų ir mašininio mokymosi gali apimti apie juos padarytas išvadas ir jų elgesio prognozes. Tačiau iš jo asmens duomenų padarytos išvados apie asmenį paprastai nėra laikomos saugotinais asmens duomenimis“ (Macmillan, 2019, p. 21) ir nors „įstatymai dažnai riboja privatumo apsaugą, kad būtų ištaisyti, blokuojami arba ištrinami asmens duomenys, kurie įvedami į algoritmus, bet ne tų duomenų vertinimas ar tokiu vertinimu pagrįsti sprendimai“ (Macmillan, 2019, p. 21). Todėl “vienas ypatingas algoritminio asmens duomenų tvarkymo iššūkis yra naujų duomenų generavimas. Kai duomenų subjektas dalijasi keletu atskirų duomenų, dažnai galima tuos duomenis sujungti, sukuriant antros ir net trečios kartos duomenis apie asmenį. Du nekenksmingi duomenys, palyginti su daug didesniu duomenų rinkiniu, gali „daugintis“ ir generuoti „kūdikio duomenis“, kurių pobūdis duomenų subjektui gali būti visiškai nenuspėjamas (Algorithms and human rights, 2017, p. 13), o tokių duomenų pagrindu mašininio mokymosi algoritmas gali padaryti išvadas ir jos gali būti panaudotos tolimesniems veiksmams generuoti, ribotai informacijai filtruoti, ypač kai kalbama apie įvairias paieškos sistemas. Duomenų panaudojimas gali turėti „įtakos asmens informaciniam apsisprendimui. Paieškos sistemos gali turėti panašų poveikį teisei į privatumą ir duomenų apsaugą, nes jos taip pat palengvina duomenų apie konkretų asmenį apibendrinimą (Algorithms and human rights, 2017, p. 13) ir asmens priskyrimą tam tikrai grupei. Tokie duomenys dažnai pavadinami **išvestiniais** *(angl. inference, inferent)* duomenimis, kur „duomenų apsaugos požiūriu pagrindinis klausimas yra tai, ar informacija turėtų būti laikoma naujais asmens duomenimis, kurie skiriasi nuo duomenų, iš kurių ji buvo padaryta“ (Sartor, *et al*, 2020, p. 39) arba išvados padarytos naudojantis asmens duomenimis turėtų būti klasifikuojamos kaip asmens duomenys. Teigiamo atsakymo atveju, tokie duomenys turėtų būti apsaugomi pagal tokias pačias taisykles, kokios taikomos asmens duomenų apsaugai ir esamas BDAR tikėtina ne iki galo sprendžia šį klausimą, nes „duomenų apsaugos įstatyme pagrindinis dėmesys skiriamas mechanizmams, skirtiems valdyti įvesties apdorojimo pusę“ (Wachter, Mittelstadt, 2019, p. 21).

Sandra Wachter ir Brent Mittelstadt teigia, kad „pats automatizuoto sprendimų priėmimo ir profiliavimo analizės nenuspėjamumas gali būti žalingas asmenims“ (Wachter, Mittelstadt, 2019, p. 20), o „brandžios informacinės visuomenės tendencija

kurti, bendrinti, parduoti ir saugoti duomenis, profilius ir kitą informaciją apie asmenis kelia papildomų iššūkių“ (Wachter, Mittelstadt, 2019, p. 20), nes yra tik keletas mechanizmų duomenų apsaugos teisėje Europoje, kurie nukreipti į išvestis (rezultatus), gautus apdorojus duomenis, tame tarpe numanomi ir išvestiniai duomenys, profiliai ir sprendimai ir jie yra apsaugomi silpnai.

Ar išvestiniai duomenys yra asmens duomenys vienareikšmiško atsakymo šiai dienai nėra, nes Europos Sąjungos Teisingumo Teisme (ESTT) nagrinėti atvejai pasižymi skirtinga praktika dėl asmens duomenų traktavimo. Yra svarbios ESTT bylos, kuriose pasiakyta dėl išvestinių duomenų traktavimo. Vienu atveju (*Jungtinė byla C–141 ir 372/12*) teismas nagrinėdamas klausimą dėl migracijos leidimo prašymo vertinimo išvados duomenų (kaip išvestinių duomenų) yra nusprendęs, ,kad „protokole (dokumente, kuriame yra nagrinėjančio pareigūno motyvai) esantys duomenys, susiję su prašytoju išduoti leidimą gyventi, ir, jei reikia, protokole esantys teisinės analizės duomenys yra „asmens duomenys“ (Wachter, Mittelstadt, 2019, p. 30), tačiau pats vertinimas negali būti klasifikuojamas kaip asmens duomenys. „Šiame sprendime nurodoma, kad tik asmens duomenys, esantys arba naudojami atliekant teisinę analizę, bet ne pati analizė, yra asmens duomenims, kuriems taikoma apsauga pagal 1995 m. Duomenų apsaugos direktyvą“ (tuo metu galiojęs teisės aktas, kurį pakeitė BDAR) (Wachter, Mittelstadt, 2019, p. 30). Ši ESTT byla yra svarbi, nes joje buvo prašoma išaiškinti, ar teisinė analizė gali būti suprantama kaip asmens duomenys. Šis apsisprendimas labai svarbus išvadų teisniam statusui, nes teisinė analizė yra „yra panaši į asmens duomenų analizę, kai gaunami arba daroma išvada“ (Wachter, Mittelstadt, 2019, p. 31).

Kitu atveju, byloje *Peter Nowak v. Data Protection Commissioner* (C-434/16) ESTT nesilaikė savo praktikos ir pasisakė, kad duomenys apie egzaminuotojo vertinimą (išvestinis duomuo, nes nusako kaip vertintojas subjektyviai pateikė savo vertinimą) yra asmens duomenys. Byloje kandidatas laikantis egzaminą (Peter Nowak) paprašė pasinaudoti teise susipažinti su vertinimu pagal savo pateiktus atsakymus. ESTT nustatė, kad tiek egzamino darbas, tiek vertintojo komentarai yra kandidato asmens duomenys, plečiamai pritaikydamas asmens duomenų apibrėžimą, kuris apima „nuomones ir vertinimus“, kurie susiję su asmeniu, t.y. duomenų subjektu. Teismo vertinimu tiek kandidato atsakymai, tiek ir egzaminuotojo įvertinimas laikomi asmens duomenimis, ir argumentavo, kad vertinimas, komentarai ir įvertinimas gali turėti poveikį kandidatui ir jo asmeniniam gyvenimui ir todėl tai yra asmens duomenys. Šis interpretavimas keistas, nes patys egzamino klausimai neturėjo jokių asmens duomenų (Wachter, Mittelstadt, 2019, p. 33).

Šios dvi bylos buvo ir yra labai svarbios BDAR suteikiamos apsaugos kontekste, nes *Peter Nowak v. Data Protection Commissioner* byla parodė, kad duomenų subjektas negali pakoreguoti kaip jis yra vertinamas, o galima tik užtikrinti, kad jo įvesti duomenys yra teisingi. „Tačiau abiem atvejais atrodo neįsivaizduojama, kad galutinis įvertinimas ir sprendimas, pavyzdžiui, sprendimas atsisakyti leisti gyventi ar neišlaikyti egzamino, nėra asmens duomenys“ (Wachter, Mittelstadt, 2019, p. 46). Kai kalbame kad automatinių sprendimų priėmimų atžvilgiu apsaugą suteikia BDAR, mašininio mokymosi ir dirbtinio intelekto technologijos atveria kitą interpretaciją, kaip traktuoti asmens duomenis, nes yra suteikiamos apribotos teisės išvestiniams duomenims. „Šiose bylose priimtuose nutarimuose ir nuomonėse išaiškinta, kad duomenų apsaugos įstatymo kompetencija nėra vertinti sprendimų ir vertinimų motyvų tikslumo ar pačių sprendimų ir vertinimų tikslumo. Atvirkščiai, reikia žiūrėti kitus įstatymus ir valdymo mechanizmus, kurie taikomi konkrečiam atvejui (pvz., apeliacijos dėl rezidentūros ar egzaminų sprendimų priėmimo procesas)“ (Wachter, Mittelstadt, 2019, p. 48), o tai reiškia, kad asmens duomenų apsaugos įstatymas negali garantuoti teisėto sprendimų priėmimo.

2.2.Asmens duomenų profiliavimo poveikis privatumui ir duomenų apsaugai

Galima teigti, kad asmens duomenų apsaugą reguliuoja BDAR ir jokios teisinės problemos šioje srityje neturėtų kilti. Algoritmų naudojimas prekyboje, finansinėse paslaugose vystosi, nes verslo galimybės pasitekti duomenis didėja, kaip didėja ir rizika. Wachter sako, kad „dėl to, kad įmonės plačiai diegia išvadinę analizę, skirtą profiliavimui, stumdymui, manipuliavimui ar automatizuotam sprendimų priėmimui, šie „privatūs sprendimai“ gali labai paveikti asmenų privatumą“ (Wachter Mittelstadt, 2019, p. 50). Dažnai „didelius duomenis ir mašininį mokymąsi įgalina tarpininkai, pavyzdžiui, trečiųjų šalių duomenų brokeriai, prekiaujantys asmens duomenimis. Asmens duomenų perdavimas sukelia pažeidimo ir tapatybės vagystės, įkyrios rinkodaros ir kitų privatumo pažeidimų riziką“ (Macmillan, 2019, p. 41). Visgi esminis klausimas, kurį reikia adresuoti mašininio mokymo algoritmų veikimui, yra tai, kaip jis elgiasi su asmens duomenimis ir kaip traktuoti naujai sukuriamus arba išvestinius duomenis apie asmenį, kokius teisinius padarinius tokie duomenys gali turėti asmens privataus gyvenimo apsaugai.

Viena pagrindinių problemų, kurios yra susietos su asmens duomenų apsaugos ir privatumo pažeidimu taikant mašininio mokymosi priemones yra profiliavimas – bet kokios formos automatizuotas asmens duomenų tvarkymas, kai asmens duomenys

naudojami siekiant įvertinti tam tikrus su fiziniu asmeniu susijusius asmeninius aspektus, visų pirma siekiant išanalizuoti ar numatyti aspektus, susijusius su to fizinio asmens darbo rezultatais, ekonomine situacija, sveikatos būkle, asmeniniais pomėgiais, interesais, patikimumu, elgesiu, buvimo vieta arba judėjimu (BDAR, 4 str. 4 p.). „Mašininiu mokymusi pagrįsti metodai, dažnai yra skirti daryti išvadas – klasifikacijas, prognozes ar sprendimus – kai jie taikomi duomenims apie asmenis“ (Sartor, *et al,* 2020, p. 39). „Profiliavimas savaime reiškia internete prieinamų duomenų ekstrapoliavimą automatizuoto informacijos rinkimo ir tolesnio profilių kūrimo bei taikymo procesuose. Profiliavimo metodai gali būti naudingi asmenims ir visuomenei, nes, pavyzdžiui, geriau skirstoma rinka arba galima atlikti rizikos ir sukčiavimo analizę (Algorithms and human rights, 2017, p. 15). Konkretūs pavyzdžiai gali būti, kad pagal asmens finansinę istoriją ir pagal jų veiklą internete gali būti vertinamas paskolos prašančio asmens kreditingumas arba pagal asmens valgymo ir fizinės veiklos įpročius jam gali būti nustatoma širdies ligų tikimybė, o tai turi įtakos jo sveikatos draudimo įmokai ar sprendimui dėl paskolos suteikimo. Mašininio mokymosi koreliacijos gali būti susijusios su asmens polinkiais ir reakcijas į dirgiklius, kai daroma teisėta ar neteisėta įtaka žmogaus elgesiui. „Pavyzdžiui, sistema sužino ryšį tarp tam tikrų asmens savybių ir veiklos (pirkinių, patiktukų *(angl. likes)* ir pan.) ir jo, kaip konkretaus vartotojo tipo, profilio ir kad sistema taip pat išmoko (arba jai buvo pasakyta), kad tokio tipo vartotojas domisi tam tikrais produktais ir gali reaguoti į tam tikro tipo skelbimus“ (Sartor, *et al,* 2020, p. 39) todėl šias savybes turinčiam asmeniui gali būti siunčiami pranešimai, kurie greičiausiai paskatins norimą prikimo elgesį“ (Sartor, *et al,* 2020, p. 40), kas susiję su vartotojų teisių apsaugos aspektais. Tokie mechanizmai daro poveikį asmens informuotam sprendimui dėl prekių ir paslaugų įsigijimo, poveikį kainai, kurią jis galėtų mokėti.

Atsižvelgiant į keliamą problematiką dėl profiliavimo duomenų, bendroji išvada doktrinoje daroma, kad profiliavimo būdu gaunami duomenys turėtų būti laikomi asmens duomenimis (minėta ESTT praktika nevienareikšmiškai traktuoja išvestinius duomenis). Tai galima pagrįsti pavyzdžiu atskiriant bendrąsias koreliacijas, kurias fiksuoja išmoktas algoritminis modelis ir tokio modelio pritaikymą konkrečiam individui paskolos paraiškos rezultatui pasiekti. Mašininio mokymosi sprendime dėl paskolos asmeniui vertinimo „mokymo rinkinį sudaro asmens duomenys: pvz., apie kiekvieną paskolos gavėją, jo vardas, pavardė, apie jį surinkti duomenys – amžius, ekonominė būklė, išsilavinimas, darbas ir kt. – ir informacija apie tai, ar jis nesumokėjo paskolos. Išmoktame algoritminiame modelyje nebėra asmens duomenų, nes jis susieja bet kokius galimų įvesties reikšmių derinius (prognozes) su atitinkama numatytosios tikimybe

(tikslu). Algoritminiame modelyje įterptos koreliacijos nėra asmens duomenys, nes jos taikomos visiems asmenims, turintiems panašių savybių. Galime juos vertinti kaip grupinius duomenis, susijusius su tokių asmenų visuma (pvz., tiems, kuriems priskirta didesnė įsipareigojimų nevykdymo tikimybė, nes jie turi mažas pajamas, gyvena skurdžioje kaimynystėje ir pan.). Tarkime, kad tada įvesties duomenims, kuriuos sudaro naujo pareiškėjo aprašymas, taikomas algoritminis modelis, siekiant nustatyti to pareiškėjo įsipareigojimų nevykdymo riziką. Šiuo atveju tiek pareiškėjo aprašymas, tiek modelio jam priskirta įsipareigojimų nevykdymo rizika yra asmens duomenys, iš kurių pirmasis yra renkami duomenys, o antrasis – išvestiniai duomenys“ (Sartor, *et al,* 2020, p. 40). Išvada, kad tais atvejais kai asmens duomenų panaudojimas algoritminiams skaičiavimams sudaro prielaidas identifikuoti asmenį ir priskirti jį tam tikrai asmenų grupei, profilis tampa išvestiniais ir turėtų būti klasifikuojami kaip asmens duomenys ir jiems turėtų būti taikomos asmens duomenų tvarkymo reguliavimo nuostatos. Kadangi asmens duomenų apsaugą reglamentuoja BDAR, asmuo tų duomenų atžvilgiu įgyja analogiškas teises, kurias turi asmens duomenims. „Teisės į skaidrumą gali padėti asmenims žinoti, kada ir kokios išvados daromos“ (Wachter, Mittelstadt, 2019, p. 51).

Iš esmės BDAR visiškai nereguliuoja duomenų priskirtų tam tiktai grupei klausimo, o tai reiškia, kad automatizuota vertinimo ir sprendimų priėmimo sistema, kuri yra pagrįsta profiliu, nors yra ir nešališka ir skirta naudingiems tikslam pasiekti, gali neigiamai paveikti atitinkamus asmenis (Sartor, *et al,* 2020, p. 23).

Apsaugą nuo profiliavimo reglamentuoja BDAR 21 straipsnis, kur asmuo turi teisę nesutikti, kad jo duomenys būtų tvarkomi rinkodaros tikslais, įskaitant ir profiliavimą. Taip pat, BDAR 22 straipsnio 1 dalis, kurioje numatyta, kad duomenų subjektas turi teisę, kad jam nebūtų taikomas tik automatizuotu duomenų tvarkymu, įskaitant profiliavimą, grindžiamas sprendimas, dėl kurio jam kyla teisinės pasekmės arba kuris jam panašiu būdu daro didelį poveikį. „Kartu šios teisės gali suteikti duomenų subjektams reikšmingą galimybę kontroliuoti išvadas, kurios gali pažeisti jų privatumą arba pakenkti jų reputacijai“ (Wachter, Mittelstadt, 2018, p. 56). Šis BDAR straipsnis turi išimčių, kai duomenų subjektas duoda sutikimą tokiam veiksmui su jo asmeniniais duomenimis (BDAR 22 straipsnis, 2 dalis, c punktas). Dirbtinio intelekto akto pasiūlyme referuojama į BDAR 22 straipsnio nuostatų įgyvendinimą, todėl būtų logiška teigti, kad problemos turėtų nebūti ir ateityje. Visgi doktrinoje nagrinėjami praktiniai klausimai, kaip ši nuostata veikia ir gali turėti įtakos asmens turtinių ir neturtinių teisių įgyvendinimui, kai mašininio mokymosi algoritmai veikia ne lokaliai, o internetinėje erdvėje, kur „sekimo ir profiliavimo programos taip pat naudojamos tikslinėje reklamoje pagal asmens

numanomų interesų profilį. Šiuo atveju vartotojo sutikimas yra svarbus reguliavimo klausimas“ (Algorithms and human rights, 2017, p. 13), kadangi duomenų rinkimo atveju vartotojas sutikdamas tiek su privatumo politika, tiek su slapukais, kurie įrašomi į naudojamus įrenginius sudaro prielaidas susilpninti galimybes ginti savo pažeistą teisę, sutartinės civilinės atsakomybės priemonėmis.

Daugelis DI sistemų priimamų sprendimų patenka į BDAR 21 straipsnio 1 dalies taikymo sritį, nes dirbtinio intelekto algoritmai vis dažniau naudojami įdarbinant, skolinant, suteikiant galimybę gauti draudimą, sveikatos paslaugų ir kt. Labiau tikėtina, kad sprendimas bus pagrįstas tik automatizuotu sprendimu, o asmenys neturės galimybės analizuoti ir peržiūrėti kaip naudojama jo informacija. Peržiūros ir paaiškinimo kaštai bus pernelyg dideli. Reiškia, kad yra rizika, jog asmuo neturės galimybės pasinaudoti savo teise, o verslas ieškos būdų kaip išvengti atsakomybės. „Duomenų subjektai ne visada gaus informaciją iš kiekvieno duomenų valdytojo, tvarkančio jų duomenis. Jei duomenis perduodantis duomenų valdytojas į pradinį duomenų subjekto atskleidimą įtraukė informaciją apie galimus trečiųjų šalių gavėjus (kategorijas), duomenų gavėjas neprivalo pateikti papildomos informacijos dėl perdavimo“ (Wachter, Mittelstadt, 2019, p. 53). Ir nors asmuo gautų išrašą apie jo tvarkomus duomenis (tokią teisę jam suteikia BDAR 15 straipsnis), jis gaus išrašą apie tvarkomas duomenų kategorijas, bet ne duomenų panaudojimo detales. Jeigu pavyktų gauti išvestinius ir numanomus duomenis, reikia pažymėti, kad BDAR 15 straipsnio 3 dalyje numatyta teisė gauti kopiją ,negali daryti neigiamo poveikio kitų teisėms ir laisvėms“.

Taigi BDAR 22 straipsnis leidžia „duomenų subjektams įvertinti ir užginčyti automatizuotus sprendimus ir profiliavimą, kurie gali būti pagrįsti išvadomis“ (Wachter, Mittelstadt, 2019, p. 78). Tokia teisė „suteikia duomenų subjektams galimybę ginčyti automatizuotus sprendimus tuose sektoriuose, kuriuose žmogiškieji sprendimai gali būti neginčytini arba kuriuose gali nebūti atitinkamų teisinių ar etinių sprendimų priėmimo standartų“ (Wachter, Mittelstadt, 2019, p. 79). Kadangi mašininio mokymosi sprendimai daugeliu atvejų yra pritaikomi pakankamai rizikinguose ir brangiuose sektoriuose, realios galimybės ginčyti tokius sprendimus gali ir nebūti.

2.3.Probleminiai teisės į privatumą ir duomenų apsaugą gynimo aspektai

Pagal esamą reguliavimą asmuo atrodytų turi visas galimybes ginti savo teisę į privatumą ir pats valdyti savo asmeninę informaciją. „Įtampa tarp profiliavimo, diskriminacijos, privatumo ir duomenų apsaugos įstatymų jau seniai pripažįstama. Šiuo atžvilgiu sąvoka

„duomenų apsauga“ yra klaidinanti, nes leidžia manyti, kad įstatymais siekiama apsaugoti duomenis, nors iš tikrųjų ja siekiama apsaugoti žmones“ (Wachter, Mittelstadt, 2019, p. 82). Tą sako ir Europos Sąjungos Teisingumo Teismas vadovaujasi savo išreikšta pozicija, kad „duomenų apsaugos įstatymas nėra skirtas užtikrinti sprendimų priėmimo procesų ar geros administracinės praktikos tikslumą“ (*Peter Nowak v. Data Protection Commissioner, C-434/16*), o tai reiškia kad BDAR nesaugo asmens privatumo, o saugo tik asmens duomenis. „Šis požiūris turi didelę reikšmę teisinei apsaugai nuo numanomų duomenų“ (Wachter, Mittelstadt, 2019, p. 58).

„Sėkmingas mašininio mokymosi modelių veikimas ir jų išvesties tikslumas priklauso nuo įvesties duomenų kokybės. Duomenų apsaugos ir privatumo įstatymai įmonėms vis dažniau nustato teisinę atsakomybę užtikrinti jų turimų ir tvarkomų duomenų tikslumą. Tačiau jie nenustato didelių duomenų ir mašininio mokymosi sistemų išvesties tikslumo“ (Macmillan, 2019, p. 4). „Tai reiškia, kad išvestiniai duomenys (vertinimai ar nuomonės) ir išvestų duomenų motyvai (net jei jie laikomi asmens duomenimis ir objektyviai neteisingi) negali būti pataisyti pagal duomenų apsaugos įstatymus ir gali būti ginčijami tik tuo atveju, jei yra nustatyta vertinimo užginčijimo procedūra“ (Wachter, Mittelstadt, 2019, p. 59). Įvairūs autoriai, nagrinėdami duomenų apsaugos ir dirbtinio intelekto santykį pabrėžia, kad nors BDAR apsaugo asmens duomenis įvairiais aspektais, tačiau išvados iš duomenų yra mažiausiai apsaugotos ir kelia bene didžiausią pavojų privatumui ir diskriminacijai. Nors asmuo pagal BDAR turi visą aibę priemonių apsaugoti savo asmens duomenis, tačiau teigiama, kad perteklinė apsauga gali apriboti privataus ir viešojo intereso galimybes vystyti inovacijas, kurių pageidauja pats vartotojas ir ekonomika, ir tada galimybės pasinaudoti teisėmis susiduria su reguliavimu kitose komerciniais interesais pagrįstose srityse, tokiose kaip intelektinė nuosavybė, komercinės paslaptys. Wachter sako, kad galimybė pasinaudoti duomenų ištrynimo teise yra paremta komerciniu interesu ir finansiniais ištekliais, todėl tuo puikiai gali pasinaudoti bendrovės. „Duomenų subjektams būtų leidžiama ištrinti tik tuos asmens duomenis, kuriuos jie pateikė, ir tik tuo atveju, jei tai neprieštarauja duomenų valdytojo verslo interesams“ (Wachter, Mittelstadt, 2019, p. 63).

Remiantis autoriais, automatinis sprendimų priėmimas vyrauja kredito, draudimo, rinkodaros srityse duomenų subjektas tikėtina susidurs su teisiniais barjerais atskirose srityse, nes atsiras reguliavimo, standartų kitų reikalavimų neapibrėžtumas, o tik „teisė ginčyti suteikia mažai apsaugos nuo automatizuotų sprendimų ir pagrindinių išvadų be tokių papildomų standartų“ (Wachter, Mittelstadt, 2019, p. 79). Taikant BDAR teisę ginčyti (t.y. panaikinti arba pakeisti automatizuotą sprendimą), bus galima įgyvendinti

sėkmingai „tik tuo atveju, jei įvesti duomenys buvo neteisingi arba neišsamūs arba bus pažeisti kiti duomenų apsaugos principai (pvz., duomenų valdytojas neįrodo teisėto duomenų tvarkymo pagrindo). Sprendimų motyvus ar parametrus galima ginčyti tik tuo atveju, jei papildomų sprendimų priėmimo standartai (pvz., kovos su diskriminacija įstatymas) egzistuoja už duomenų apsaugos įstatymo ribų, kurie patys nenustato sprendimų priėmimo procesų turinio ar rezultatų standartų“ (Wachter, Mittelstadt, 2019, p. 80). Taigi galima daryti išvadą, kad išvestiniai ir numanomi duomenys, pagal duomenų apsaugos įstatymą gauna mažiau apsaugos, nei asmens duomenys pateikti pačio asmens.

„Dabartiniai duomenų apsaugos teisės ribojimai gali pakenkti platesnio pobūdžio tikslui apsaugoti privatumą nuo naujų technologijų keliamos rizikos“ (Wachter, Mittelstadt, 2019, p. 82) ir siūloma svarstyti naujus būdus kaip apsaugoti asmenis nuo išvestinių duomenų (didžiųjų duomenų analizės kontekste) poveikio jų privatumui, t.y. įtvirtinti teisę į pagrįstas išvadas (*angl. the right to reasonable inferences*) duomenims, kurie:

1) pažeidžia privatumą arba kenkia reputacijai arba yra didelė tikimybė, kad taip bus ateityje,
2) turi mažai patikrinamumo, nes yra nuspėjami arba pagrįsti nuomone, o naudojami svarbiems sprendimams priimti.

Teisė į pagrįstas išvadas nėra skirta automatinių sprendimų savarankiškumui pažeisti, o suteikti duomenų subjektui galimybę daugiau sužinoti apie duomenų valdytojo suvokimą ir sprendimų priėmimo procesus bei galimai įtikinti duomenų valdytoją, kad vienas ar abu klysta (Wachter, Mittelstadt, 2019, p. 99). Aišku tokia teisė sukeltų nemažai problemų kitose šakose, tokiose kaip intelektinės nuosavybės apsauga, komercines paslaptis ar net trečiųjų šalių privatumą, todėl toks pasiūlymas turi būti labai rizikingas ir šiuo metu nėra įgyvendinamas.

Analizuota problematika sako, kad ne patys asmens duomenys, bet mašininio mokymosi algoritmų elgesys su duomenimis ir po to daromos išvados kelia daugiausia klausimų. Kaip minėta anksčiau, asmens privatumas yra saugomas, ir „deliktinė atsakomybė taikoma kai atitinkami duomenys yra asmens informacija, kuri saugoma taisyklių“ (Benhamou, Ferland, 2020, p. 19). BDAR 82 straipsnis imperatyviai nustato, kad bet kuris asmuo, patyręs materialinę ar nematerialinę žalą dėl reglamento pažeidimo, turi teisę iš duomenų valdytojo arba duomenų tvarkytojo gauti kompensaciją už patirtą žalą. Norint turėti teisę teisme pareikšti ieškinį dėl kompensacijos išieškojimo, paprastai reikia pareikšti, kad buvo padaryta žala. Iš keliamų bylų dėl žalos atlyginimo įvairių jurisdikcijų „teismams buvo sunku nustatyti žalą dėl duomenų apsaugos ir privatumo

įstatymų pažeidimų, todėl teisiniai požiūriai dažnai skiriasi. Daugelis skundų buvo atmesti, nes vartotojai neįrodė patirtos žalos“ (Macmillan, 2019, p. 56).

Atitinkamai galima panagrinėti tikslinės reklamos sritį, kur „atrodo, kad nėra nieko blogo, jei vartotojams pateikiami jų pomėgius atitinkantys skelbimai, padedantys jiems naršyti tarp daugybės internete siūlomų parinkčių. Tačiau asmeniniams poreikiams pritaikyta reklama apima masinį asmens duomenų rinkimą, kuris naudojamas reklamuotojų ir tarpininkų interesais, galbūt prieš duomenų subjektų interesus. Tokie duomenys iš tiesų suteikia naujų įtakos ir kontrolės galimybių, jie gali būti panaudoti apgaulingoms, agresyvioms žinutėms arba apskritai žinutėms, kurios apeina racionalumą, apeliuojant į silpnybes ir emocijas (Sartor, *et al*, 2020, p. 25). Toks reklamos elgesys gali paskatinti asmenis sudaryti rizikingus finansinius sandorius, įsigyti jiems nereikalingų prekių ir taip sukelti materialią ir nematerialią žalą, nors įrodyti, kad asmuo buvo paveiktas reklamos ir juo buvo manipuliuojama bus labai sunku (arba beveik neįmanoma). Iš to kyla atsakomybės problematika, nes atsakomybės „priklauso nuo atsakomybės pagal įstatymus, įskaitant padarytos žalos atlyginimą. Vienas iš sunkumų kuriant politiką, teisinius įsipareigojimus ir vartotojų teisių gynimo priemones duomenų apsaugos srityje kyla dėl nematerialaus žalos, nuo kurios vartotojas turi būti apsaugotas arba už kurią jiems turi būti atlyginta, pobūdžio“ (Macmillan, 2019, p. 55). Skirtingi autoriai pastebi, kad užginčyti automatizuoto sprendimų priėmimo nesąžiningumą yra labai sunku. „Tiesą sakant, mašininio mokymosi sistemų prognozės yra pagrįstos statistinėmis koreliacijomis, dėl kurių gali būti sunku ginčytis remiantis individualiomis aplinkybėmis, net jei būtų pateisinamos išimtys“ (Sartor, *et al*, 2020, p. 21), be to, asmens duomenys nėra laikomi turtinių santykių objektu, o neturtinės žalos įrodinėjimas galimai bus brangiai kainuojantis.

Todėl svarbus yra atskyrimas ir išgryninimas mašininio mokymosi algoritmų veikimo ir jų įtakos asmens privatumui, reputacijai ir su juo susijusių teisėtų sprendimų priėmimo, ir teisėtos asmens duomenų apsaugos reikalavimų, kurie taikomi konkretaus asmens duomenims. Daugumoje atvejų, mašininio mokymosi sprendimų priėmimas asmens atžvilgiu nėra asmens duomenų reguliavimo pažeidimas, todėl skirtingoms sritims taikomų reguliacinių reikalavimų derinimas turi užtikrinti pamatines asmens teises į privatų gyvenimą.

III. MAŠININIO MOKYMOSI SPRENDIMŲ POVEIKIS VARTOTOJŲ TEISĖMS

Vartotojų teisių apsaugos klausimas mašininio mokymosi sprendimų atžvilgiu daugiausia dėmesio sulaukė, kai įsigalėjo elektroninė prekyba ir klasikinis pirkėjo - pardavėjo teisinis santykis persikėlė į skaitmeninę erdvę. „Vartotojų apsauga formuluojama įvairiai, tačiau dažniausiai siekiama propaguoti sąžiningumo, atskaitomybės ir skaidrumo vertybes. Politinės diskusijos apie vartotojų apsaugos ryšį su dirbtiniu intelektu ir mašininiu mokymusi yra susijusios su algoritmų ir mašininio mokymosi sistemų gebėjimu atspindėti tokias vertybes“ (Macmillan, 2019, p. 19).

Europos Komisija 2018 metais pateikė pasiūlymą peržiūrėti reguliavimą dėl ekonominių vartotojų interesų apsaugos paskelbus komunikatą „Naujos galimybės vartotojams“, motyvuodama pokyčiais, kad įmonių santykiai su vartotojais būtų sąžiningi ir skaidrūs (2018 m. balandžio 11 dienos Europos komisijos pasiūlymas). „Dėl daugelio priežasčių vartotojai gali būti pažeidžiami, kai naudojasi paslaugomis, kurios priklauso nuo kompiuterinio apdorojimo. Jų veikimas pranoksta daugumos gyventojų supratimą. Dėl to kompiuteriai ir jų sukurti rezultatai gali būti suvokiami kaip objektyvūs ir netgi teisingi. Tačiau šiandien yra rizika, kad kai kurie skaitmeninių paslaugų aspektai vartotojams atrodys nesąžiningi, neatskaitingi ir neskaidrūs, o tai sumažins vartotojų ir paslaugų teikėjų pasitikėjimą ir taip trukdys skaitmeninių paslaugų augimo perspektyvoms.“ (Macmillan, 2019, p. 19). Ekonomikos požiūriu prekiautojų elgesys vartotojų atžvilgiu gali daryti didelį poveikį vartotojų rinkų veikimui. Taip yra todėl, kad tokiose rinkose prekiautojai turi labai didelę įtaką vartotojų informavimui ir sprendimų priėmimui (2018 m. balandžio 11 dienos Europos komisijos pasiūlymas).

Šioje darbo dalyje aptariamos pagrindinės su vartotojų apsauga susijusios mašininio mokymosi problemos ir kaip Omnibus direktyva prisidės sprendžiant vartotojų teisių apsaugos klausimus.

3.1.Algoritmų poveikis vartotojų teisėms

Duomenys surinkti apie asmenį ir išvestiniai duomenys suteikia paslaugų teikėjams daug galimybių vystyti savo verslą ir pasiekti rezultatus. Įmonės gali naudoti surinktus duomenis, kad padarytų išvadą apie vartotojų norą mokėti. „Didelių duomenų prieinamumas palengvina kainų diskriminaciją“ (OECD, 2018, p. 2), finansinių paslaugų srityje duomenų prieinamumas leidžia geriau įvertinti vartotojo keliamą riziką ir pasiūlyti

paslaugas, kurių kitu atveju nebūtų galima gauti. „Tačiau galimybė turėti daug duomenų apie vartotoją taip pat sukuria informacijos asimetriją, kai paslaugų teikėjas apie vartotoją žino daugiau, nei vartotojas apie tiekėją“ (Macmillan, 2019, p. 40). Paslaugų tiekėjas gali pasinaudoti šia situacija ir pradėti taikyti individualizuotą kainodarą, „kai paslaugų teikėjas skirtingiems vartotojams už tą patį produktą taiko skirtingas kainas“ (Macmillan, 2019, p. 40). Kartais diferencijuota kainodara yra naudinga vartotojams, pavyzdžiui taikant skirtingas kainas studentams ar vyresnio amžiaus žmonėms pritaikant nuolaidą. Tačiau nesąžiningumas gali atsirasti tais atvejais kai tam tikros gyventojų grupės moka didesnę kainą atsižvelgiant į jų profilį, dėl geografinės padėties ar kitų savybių. Tarkime, finansinių paslaugų srityje skirtinga kainodara pirmiausia yra susijusi su vartotojo rizikos profiliu, taip padidinant ekonominį efektyvumą ir atgrasant asmenis nuo neteisėtos veiklos. Draudimo rinkoje tai svarbus elementas, kai didesnės rizikos asmenys gauna mažiau palankią kainodarą. „Tačiau skirtinga draudimo produktų kainodara gali sukelti nesąžiningumą, kai atsiranda rizikos veiksnių, kurių asmuo negali kontroliuoti, pvz., sveikatos draudime“ (Macmillan, 2019, p. 40). Algoritmų pagalba „dideli duomenys gali siūlyti skirtingą kainodarą, iš asmens duomenų darydami išvadas apie asmens poreikį gauti paslaugą, jo gebėjimą mokėti ir jautrumą kainai“ (Macmillan, 2019, p. 40), kai kaina parenkama maksimaliai artima sumai, kurią profiliuotas vartotojas gali mokėti. Kadangi vartotojas nežino visos informacijos apie paslaugos teikėją, tai negali susitarti dėl minimalios kainos, kurią tiekėjas norėtų gauti, kad jam atsipirktų investicijų grąža. Tokia situacija sukelia sutarties laisvės principo pažeidimo implikacijų.

OECD 2018 metų pranešime „Individualizuotos kainodaros reguliavimas skaitmeninėje eroje“ (*angl. The regulation of personalized pricing in digital era)* teigė, kad jokių empirinių duomenų, kad internetinėje rinkoje egzistuoja individualizuota kainodara nėra, tačiau tam tikri atvejai rinkoje rodo, kad vartotojai susiduria su atvejais, kai „ištrynus slapukus iš savo kompiuterio naršyklės, jis gavo mažesnę kainą už konkretų DVD Amazon.com“ (OECD, 2018, p. 3). Vartotojų reakcijos baimė yra galimas paaiškinimas, kodėl tikslinės kainodaros beveik nesilaikoma. „Tačiau yra subtilesnių – ir vartotojų požiūriu priimtinesnių – būdų, kaip įmonė gali pasiekti tą patį rezultatą“ (OECD, 2018, p. 3), kai „įmonė gali įsitraukti į paieškos diskriminaciją arba valdymą, kurį sudaro skirtingų produktų rodymas skirtingų grupių klientams, remiantis turima informacija apie vartotojus. Pavyzdžiui, „Wall Street Journal“ (2012 m.) pranešė, kad kelionių agentūra „OrbitzWorldwide“ rodė brangesnius viešbučių pasiūlymus „Mac“ vartotojams nei kitų kompiuterių vartotojams. Panašią praktiką taikė ir Staples.com: tas

pats laikraščio straipsnis atskleidė, kad šioje svetainėje buvo rodomos skirtingos kainos, kai buvo nustatytos potencialių pirkėjų vietos“ (OECD, 2018, p. 3).

Pranešime OECD teigia, kad turi būti nustatyti teisiniai instrumentai, kurie leistų sureguliuoti individualizuotos kainodaros klausimą ir šios sritys turi būti: vartotojų apsaugos įstatymai, duomenų apsaugos įstatymai, konkurencijos apsaugos įstatymai ir nediskriminavimo įstatymai.

Apibendrintoje lentelėje pateikiami ir palyginami pagrindiniai kiekvienos teisinės priemonės tikslai ir taikymo sritis, taip pat pasekmės naudotojų ir vartotojų skaidrumui ir pasirinkimo galimybėms bei kai kurių individualizavimo atvejų draudimui, kai manoma, kad įgaliojimų nepakanka.

2 lentelė. **Pagrindinių taisyklių, taikomų individualizuotai kainodarai, sąlygos ir poveikis**

Teisinis instrumentas	Vartotojų apsauga	Asmens duomenų apsauga	Konkurencijos apsauga	Nediskriminavimas
Pagrindiniai tikslai	Skaidrumas ir sąžiningumas atliekant sandorius	Duomenų subjekto privatumas ir apsisprendimas	Ekonominė gerovė: visa arba vartotojo	Kai kurių pagrindinių teisių apsauga
Apimtis	Daugiausia B2C sandoriai	Sandoriai, pagrįsti asmens duomenimis	Įmonių sudaryti sandoriai	Visi valstybės ir (arba) privačių įmonių sandoriai
Skaidrumas	Skirtingi skaidrumo laipsniai - vien faktas, kad kainos yra individualizuotos - personalizavimo parametrai - vidutinė arba vidutinė siūloma kaina	- Asmens duomenų naudojimas - individualios kainodaros logika		
Vartotojo pasirinkimas	Individuali kainodara gali būti laikoma nesąžininga tam tikromis aplinkybėmis, turinčiomis įtakos vartotojo pasirinkimo kokybei	Norint pagrįsti individualią kainodarą asmens duomenimis, reikalingas duomenų subjekto sutikimas	Gali paskatinti konkurenciją, taigi ir vartotojų pasirinkimą	
Suasmeninimo draudimas		Individuali kainodara pagrįsta jautriais duomenimis, paprastai yra draudžiama	Individuali kainodara, kuri kenkia gerovei (turi būti nustatoma kiekvienu konkrečiu atveju) Susitarimai - Horizontaliai - Vertikalus vienašalis elgesys - Išskirtinis B2B: vidinė ir išorinė diskriminacija - Išnaudojamas B2C	Individuali kainodara pagrįsta jautriais pagrindais, pvz., lytimi arba rasine / etnine kilme ES, individuali kainodara pagal pilietybę arba gyvenamąją vietą

Šaltinis: OECD (2018). *The regulation of personalized pricing in the Digital era* – Note by Marc Bourreau and Alexandre de Strell. DAF/COMP/WD (2018)150

Šių sričių pakeitimus įgyvendina ir nauja Europos Parlemento ir Tarybos direktyvos (ES) 2019/2161 direktyvos (Omnibus direktyva), kurios nuostatos bus aptartos toliau, tačiau OECD pranešime yra pastebėta, kad „nors dėl sparčios duomenų rinkimo ir duomenų analizės techninės pažangos kainų individualizavimas tampa lengvesnis ir pigesnis, keli naujausi tyrimai visame pasaulyje rodo, kad skaitmeninės įmonės paprastai nesuasmenina savo kainų. Tai gali būti paaiškinta vartotojų nepasitikėjimu individualizuotomis kainomis ir vartotojų labiau priimtinomis alternatyvomis, tokiomis kaip paieškos rezultatų personalizavimas ar nuolaidos“ (OECD, 2018, p. 11).

Doktrinoje autoriai rašo, kad didžiausia profiliuoto vartotojo apsaugos rizika kyla dėl skirtingos kainodaros ir galimybės padaryti informuotą pasirinkimą. Skirtinga kainodara gali tapti diskriminacine, kai kainos yra nustatomos pagal kriterijus, kurie nors ir atrodo objektyvūs, tačiau daro neigiamą įtaką atskiroms grupėms. Toks pavyzdys gali būti, kad aukštesnė paslaugos kaina yra nustatoma asmenims priskirtiems grupei iš geografinės srities, kur stebimas istoriškai didesnis įsipareigojimų nevykdymo lygis nei kitam šalia esančiam geografiniam vienetui. Dėl požymių, kurie istoriškai buvo nepalankūs, asmenys gali būti diskriminuojami nors tai neturi įtakos jų kreditingumui. „Pavyzdžiui, asmuo, turintis gerą atlyginimą ir mažai skolingas, gali būti traktuojamas neigiamai dėl to, kad gyvena bendruomenėje (arba turi draugų socialiniuose tinkluose, ar tą patį gydytoją, ar apsiperka nuolaidų parduotuvėse), kur žmonės istoriškai turi didesnes skolas“ (Macmillan, 2019, p. 41). Taigi vartotojas, negalėdamas kontroliuoti savo duomenų ir rezultato, susiduria su diskriminacija kainos atžvilgiu, o dėl informacijos trūkumo negali (ar neturi galimybės) ginti savo teisės į teisingą kainą. Be abejo reguliavimas numato vartotojo teisę į paaiškinimą, kaip priimamas automatizuotas sprendimas, tačiau praktikoje susiduriama su dviem problemomis:

1. technologinius sprendimus yra sunku paaiškinti vartotojui jam suprantama kalba, nes (kaip minėta anksčiau) mašininio mokymosi algoritmai pasižymi neskaidrumu ir vadinami „juodąja dėže“ ir kompiuterio kodas neatskleidžia kaip sprendimas buvo priimtas;
2. dėl savo autentiškumo mašininio mokymosi algoritmų modeliai dažnai yra komercinių paslapčių ir programinės įrangos autorių teisių objektas, kuris egzistuoja konkurencingoje komercinėje rinkoje. „Mašininio mokymosi operatorius gali nenorėti dalytis mašininio mokymosi algoritmo kodavimu arba paaiškinimu, kad tai nesusilpnintų konkurencijos galimybių ir nepakenktų pradinėms investicijoms“ (Macmillan, 2019, p. 50).

Šios problemos bene analogiškos asmens duomenų apsaugos srityje, nes vartotojų apsauga yra visiškai susijusi su asmens pasirinkimo teise pačiam kontroliuoti savo elgesį, todėl bet kokie būdai paveikti vartotojo pasirinkimą jį atakuojant specialiu profiliuotu turiniu tampa priežastimi pergalvoti rinkos taisykles. Tikėtina, kad nemažai klausimų yra susiję su situacija, kad asmens duomenys nebuvo (ir daugeliu atveju nėra) laikomi turtu, kuris leistų įvertinti asmens patiriamą žalą ar nuostolius, dėl asmens pažeistos teisės. Šį asmens duomenų aspektą paliečia naujoji Omnibus direktyva, kur į vartotojų teisių apsaugos reguliavimą įtraukiamas atsiskaitymas už teikiamas paslaugas (nesvarbu naudojasi asmuo ar nesinaudoja) asmens duomenimis.

3.2.Vartotojų teisų apsaugos pokyčiai Omnibus direktyvos kontekste

Verslo ir vartotojo informacijos asimetrija yra gerai žinoma. Daug informacijos sužinoma teikiant paslaugas, kuri teikiant pasaugas vartotojams eina dviem kryptimis: „iš teikėjo vartotojui, taip pat iš vartotojo paslaugų teikėjui. Kompiuterinės sistemos, kurias valdo tiekėjai ar prekybininkai, gali stebėti, tikrinti ir analizuoti bet kurį operacijos aspektą, fiksuodamos kiekvieną klaviatūra įvedamą simbolį ir kiekvieną spustelėtą nuorodą“ (Grochowski, *et al*, 2021, p. 46). Nors kai kurios paslaugos nėra apmokamos vartotojo, tačiau tokių paslaugų grąža tiekėjui ateina per rinkodarą, „per asmeninės informacijos rinkimo ir šios informacijos panaudojimo praktiką, kuriant komunikaciją, kuri optimaliai atspindi asmenų savybes ir tuo jiems daro įtaką“ (Grochowski, *et al*, 2021, p. 46). Vartotojas dažnai nežino, kokį poveikį paslaugų tiekėjų veikla daro jo asmeniniam gyvenimui. „Tokios padėties priežastys bent iš dalies yra susijusios su neskaidrumo problema“ (Grochowski, *et al*, 2021, p. 47). „Apskritai, visos mašininio mokymosi sistemos, naudojamos vartotojų santykiams valdyti, yra neskaidrios vartotojams, kurie neturi realios galimybės apžiūrėti sistemos ir jos mokymo rinkinio, ištirti sistemos modelyje esančius prognozuotojus ar gauti paaiškinimų, kodėl atsiranda tam tikras rezultatas“ (Grochowski, *et al*, 2021, p. 48).

Kai yra kuriamas reguliavimas, Europos Sąjungoje siekiama visų pirma jį kurti taip, kad neprofesionalus vartotojas galėtų padaryti informuotą sprendimą. Dauguma skaitmeninių technologijų rinkos pokyčių reguliavimo iniciatyvų paremiamos pareiga paslaugos tiekėjui suteikti informaciją ir galimybę paaiškinti, o vartotojui suteikia teisę būti informuotam. „Informavimo taisyklių svarba sprendžiant neskaidrios algoritminės praktikos iššūkius buvo pabrėžta neseniai vykdant ES vartotojų apsaugos direktyvos įstatymo reformą“ (Grochowski, *et al*, 2021, p. 51). Tokie „teisės aktų pokyčiai yra

bandymas užkirsti kelią technologijų naudojimui vykdant internetinius sandorius vartotojų sąskaita“ (Galli, 2021, p. 53).

ES nuo 2022 m. gegužės 28 d. įsigalios nauja 2019 m. lapkričio 27 d. Europos Parlemento ir Tarybos direktyva (ES) 2019/2161 (Omnibus direktyva), kur ES “įtraukė daugybę nuostatų, paaiškinančių ir išplečiančių galiojančių normų taikymą skaitmeniniame kontekste“ (Grochowski, *et al*, 2021, p. 51). Atsižvelgiant į besikeičiančią aplinką naujas sureguliavimas keičia verslo ir vartotojų santykius, suteikiant daugiau skaidrumo vartotojams. Omnibus direktyva keičia Nesąžiningos komercinės veiklos direktyvą 2005/29/EB, Vartotojų teisių direktyvą 2011/83/ES, Nesąžiningų sutarčių sąlygų direktyva 93/13/EEB ir Kainų žymėjimo direktyvą 98/6/EB. Pakeitimų paketas leis daugiau apsaugoti vartotojus nuo nesąžiningos veiklos atsižvelgiant į skaitmeninės srities pokyčius. „Ši informacija apima pagrindinius parametrus, pagal kuriuos nustatomas vartotojams pateikiamų produktų reitingas, ir santykinę tokių parametrų svarbą“ (Grochowski, *et al*, 2021, p. 51).

Teisės doktrinoje vartotojų teisės apima tokias teises:

- teisės iki pirkimo (išankstinis įsipareigojimas), pavyzdžiui, informacija apie teikiamą prekę ar paslaugą;
- paties produkto ar paslaugos teikimas, kokybė ir funkcionavimas (įsipareigojimas);
- priemonės, skirtos paslaugų teikėjų atskaitomybei po pirkimo (po įsipareigojimo) (Macmillan, 2019, p. 19).

„Daugelio šalių įstatymai apsaugo vartotojus nuo klaidinančių gaminių aprašymų, nesąžiningų sutarties sąlygų (pvz., atsakomybės atmetimo), nekokybiškų gaminių ir žalos atlyginimo mechanizmų trūkumo. Tokie įstatymai draudžia gamintojams ir mažmenininkams derėtis dėl tokių sąlygų su vartotojais, kad jie negalėtų ginčytis, kad vartotojai sutiko su jomis pirkdami produktą ar paslaugą. Taikant vartotojų apsaugos metodą nustatomi minimalūs bendri standartai ir procedūros, kad būtų užtikrintas pagrindinis apsaugos lygis, o ne viskas paliekama vartotojų savarankiškumui ir atsakomybei“ (Macmillan, 2019, p. 20).

Omnibus direktyva keičia vartotojų teisę į individualias teisų gynimo priemones, kai jie nukenčia nuo nesąžiningos komercinės praktikos, tokios kaip agresyvi rinkodara (2018 m. balandžio 11 dienos Europos komisijos pasiūlymas). Agresyviai rinkodarai gali būti naudojamos aptartos mašininio mokymosi priemones profiliuojant asmenis pagal jų elgesį ir pomėgius ir taip pritaikant turinį, skatinant vartotoją įsigyti prekes ar paslaugas. Taip pat, elektroninėse prekyvietėse siekiama užtikrinti skaidrumą, kad vartotojas žinotų

kaip jiems yra rikiuojami pasiūlymai ir iš ko jie perka prekes ar paslaugas. Atsiranda reikalavimas identifikuoti, su kuo pirkėjas sudaro sutartį elektroninėje prekyvietėje, kad būtų lengviau nustatyti kas yra atsakingas už nesklandumus.

Vienas esminių pakeitimų – išplečiama vartotojų apsauga skaitmeninių paslaugų srityje (išplečiant Vartotojų teisių direktyvos 2011/83/ES taikymą), už kurias vartotojai nemoka pinigais, tačiau pateikia asmens duomenis (pavyzdžiui el. paštas, socialinė žiniasklaida, debesijos saugyklos paslaugos). Šis pakeitimas ypač aktualus vertinant anksčiau aptartą asmens duomenų apsaugos ir privatumo klausimą, nes asmens duomenys tampa atlygiu skaitmeninėms platformoms už teikiamas paslaugas. Skaitmeninis turinys ir skaitmeninės paslaugos dažnai teikiami internetu pagal sutartis, pagal kurias vartotojas nemoka kainos, tačiau pateikia prekiautojui asmens duomenis (2018 m. balandžio 11 dienos Europos komisijos pasiūlymas). Asmens duomenys įgauna prielaidas būti laikomi turtinių santykių objektu ir cirkuliuoti civilinėje apyvartoje, nes pagal Lietuvos Respublikos civilinis kodeksą (CK) civilinės teisės objektais laikomi daiktai, pinigai ir vertybiniai popieriai, kitas turtas bei turtinės teisės, intelektinės veiklos rezultatai, informacija, veiksmai ir veiksmų rezultatai, taip pat kitos turtinės ir neturtinės vertybės (CK 1.97 straipsnio 1 dalis). „Civilinė teisė saugo ir tokias vertybes kaip vardas, gyvybė, sveikata, žmogaus privatus gyvenimas, juridinio asmens pavadinimas. Iš esmės civilinių teisių objektai yra visa tai, su kuo yra susijusios civilinių teisinių santykių subjektų teisės ir pareigos, todėl neabejotina, jog duomenis taip pat galima laikyti civilinės teisės objektu“ (Medvedevaitė, *et al*, 2021). Šis klausimas neabejotinai taps vienas sudėtingiausių sprendžiant mašininio mokymosi sprendimais sukeltos žalos ar nuostolių atlyginimo klausimus.

Ominibus direktyva kelia reikalavimus produktų reitingavimo algoritmams ir elektroninės prekybos subjektai turi informuoti vartotojus apie pagrindinius kriterijus, kurie lemia vartotojui nurodomų produktų reitingą, pateikiamą pagal jo paieškos užklausą, kai vartotojas turi galimybę ieškoti produktų. „Pavyzdžiui, gali būti nurodoma, kad produktai yra reitinguojami pagal kainą, produktą įsigijusių pirkėjų įvertinimus ar kelių skirtingų kriterijų derinį. Jei produkto reitingas paieškos rezultatuose grindžiamas apmokėta reklama ar kitais mokėjimais, gautais iš atitinkamų prekybininkų, tokia informacija taip pat turi būti aiškiai nurodyta vartotojams“ (Lazauskaitė, 2021). „Reitingavimas reiškia prekybininkų pasiūlymų svarbą ir paieškos rezultatų svarbą, kaip jie pateikiami, sutvarkyti ir perduodami vartotojams (pavyzdžiui, dėl algoritminės sekos, vaizdinių akcentų ir vertinimo ar peržiūros mechanizmų). Reitingavimas pagal savo pobūdį laikomas vienu efektyviausių vartotojų pasirinkimo postūmių: jis neslepia visų

svarbių variantų, tačiau paieškos rezultato pozicionavimas labai sąlygoja pasirinkimo tikimybę“ (Galli, 2021, p. 53).

Pagal Ominibus reikalavimus „prekiautojams rekomenduojama identifikuoti pagrindinius kriterijus, kuriais remiantis vartotojui bus pateikiamas produktų reitingas pagal jo atliktą paieškos užklausą. Informaciją apie šiuos kriterijus aiškiai ir skaidriai pateikti vartotojui atskiroje interneto svetainės dalyje, į kurią galima tiesiogiai ir lengvai patekti iš puslapio, kuriame pateikti vartotojo užklausos rezultatai“ (Lazauskaitė, 2021). Vartotojas informaciją turi rasti ir su ja susipažinti. Gali kilti iššūkių dėl algoritmų, naudojamų reitingavimui intelektinės apsaugos, tačiau Omnibus direktyvoje „pareiga pateikti informaciją apie pagrindinius kriterijus, pagal kuriuos išrikiuojami paieškos rezultatai, nepažeidžia jokių komercinių paslapčių, susijusių su pagrindiniais algoritmais“ (2018 m. balandžio 11 dienos Europos komisijos pasiūlymas), nes turi būti pateikiamas pagrindinių standartinių kriterijų paaiškinimas, tačiau informacija neturi būti individualizuota pagal kiekvieną paieškos užklausą.

Kitas pakeitimas, turinis įtakos mašininio mokymosi sprendimų naudojimui prekyboje, yra reikalavimai dėl kainos individualizavimo taikant algoritmus. „Į Direktyvą 2011/83/ES dėl vartotojų teisių buvo įtraukta papildoma informavimo pareiga, pagal kurią prekybininkai turi informuoti vartotojus, ar kaina buvo pritaikyta individualiai, remiantis automatizuotu sprendimų priėmimu“ (Grochowski, *et al*, 2021, p. 52). Algoritmų taikymas, vartotojo perkamajai galiai nustatyti, nėra ribojamas, tačiau Omnibus direktyva įtvirtina pareigą informuoti vartotoją, kad nurodoma prekės ar paslaugos kaina gali būti padidinta būtent jam, o „verslininkas, taikantis atitinkamus algoritmus, privalo vartotoją kiekvieną kartą informuoti, kad kaina buvo individualizuota taikant automatizuotą sprendimų priėmimą“ (Lazauskaitė, 2021).

Kadangi Ominbus direktyva dar tik pradės veikti, jos poveikio šiandien pasakyti negalima, tačiau jau dabar autoriai, nagrinėjantys vartotojų apsaugos temą (Galli, Grochnowki ir kt.) pateikia tam tikrų rizikų dėl numatomų reikalavimų įgyvendinimo. Galli sako, kad pokyčiai reguliavime suteiks tik dalinę vartotojų apsaugą, nes pirma, jie dauguma paremti informavimo pagrindais ir „nedraudžiama naudoti algoritmines sistemas siekiant paveikti vartotojus reitinguojant arba individualizuojant kainas, jei tai aiškiai nurodoma vartotojui. Be to, reitingavimas ir personalizavimas yra tik du būdai, kuriuos šiuo metu naudoja internetinių platformų operatoriai.“ (Galli, 2021, p. 53), nors poveikio priemonių daryti įtaką vartotojas yra daugiau. Tai pat Galli sako, kad „naujausi teisės aktų pokyčiai neužima rimtos pozicijos dėl vartotojų sprendimų priėmimo problemų, susijusių su dirbtinio intelekto įrankių, o ne tik automatizuotų sprendimų

priėmimo sistemų, diegimu rinkodaroje“ (Galli, 2021, p. 53), nes dirbtinio intelekto technologijos sugebės pranokti žmones intelektualiais ir reaktyviais gebėjimais suprasti vartotojus.

Grochowski taip pat akcentuoja, kad dauguma apsaugos aspektų yra susiejama su skaidrumo reikalavimų nustatymu, ir išankstiniu vartotojo informavimu, tačiau mašininio mokymosi algoritmų skaidrumas nereiškia tik informavimo, o informacijos suteikimo prasmė būtų kai „vartotojai turėtų būti toliau informuojami apie įvesties duomenis, į kuriuos DI sistema atsižvelgia (pvz., dėl paskolos paraiškos: pareiškėjo pajamos, lytis, turtas, darbas ir kt.), ir apie tai, ar tokie duomenys palankūs ar nepalankūs galimiems rezultatams“ (Grochowski, *et al*, 2021, p. 56).

Apibendrinant, reikia pažymėti, kad Omnibus direktyva ir kartu keičiami kiti susieti teisės aktai yra gera pradžia reguliuoti mašininio mokymosi algoritmais priimamų sprendimų skaidrumą. Vienas esminių pokyčių, yra atskaitymo už paslaugas asmens duomenimis įtvirtinimas, tokių būdu įtvirtinant naujas galimybės sutartinių santykių institutui, o asmens duomenų atžvilgiu naujas diskusijas dėl asmens duomenų, kaip turtinių santykių ir civilinės apyvartos objekto apibrėžties. Reikalavimai dėl informacijos atskleidimo vartotojui kelia nemažai susirūpinimo, ar pats vartotojas iš tiesų galės pasinaudoti suteikiamomis teisėmis į informavimą ir, ar to pakaks? Tai, kad toks reguliavimas atrodo per silpnas vartotojų teisių gynimo požiūriu, tikėtina turi būti atsverta bendrąją inovacijų skatinimo politika, nes per didelis spaudimas iššauks rizikas vystyti naujus inovatyvius sprendimus ir technologijas, kurios turi atitikti laikmetį. Neišvengiamai įsigaliojus naujam reguliavimui iškils praktinių iššūkių vartotojų ir paslaugų tiekėjų teisinių santykių ir konfliktų sprendimo srityje, kuriuos teismams teks spręsti didelio neapibrėžtumo sąlygomis.

IV. MAŠININIO MOKYMOSI SPRENDIMŲ KELIAMOS ATSAKOMYBĖS PROBLEMOS

Darant skirtingų sričių analizę, mašininio mokymo sprendimai kelia svarbius atsakomybės klausimus, kurie turi būti pritaikomi prie esamo reguliavimo. „Paprastai atsakomybę užtraukia žala esminiams asmens interesams, tokiems kaip gyvybė, sveikata, kūno neliečiamybė, judėjimo laisvė, privati nuosavybė, o kai kuriose šalyse ir grynai ekonominiai nuostoliai bei žala žmogaus orumui“ (Montagnani, Cavallo, 2021, p. 214). Neabejotinai tai yra vienas iš esminių mašininio mokymo ar kitų dirbtinio intelekto sprendimų keliamų klausimų. Šie klausimai kyla, nes kaip sako Reed yra akivaizdu, kad kai kurios mašininio mokymosi technologijos gali sukelti fizinius sužalojimus, arba sugadinti kitą turtą, pavyzdžiui transporto priemones kai yra savarankiškas vairavimas (be vairuotojo). Mašininis mokymasis taip pat naudojamas medicininei diagnostikai ir gydymui, kur yra taip pat aiškus fizinio sužalojimo pavojus, o atsakomybės klausimą apsunkina neaiškumai, kokį našumą turėtų pasiekti mašininio mokymosi technologija. Nagrinėjant mašininio mokymosi keliamus žmogaus teisių pažeidimo klausimus, taip pat atkreiptinas dėmesys, kad „nemateriali žala taip pat gali būti padaryta, jei mašininio mokymosi technologija neteisingai atskleidžia tam tikros rūšies informaciją. Toks atskleidimas gali sukelti atsakomybę pagal konfidencialumo įstatymą arba duomenų apsaugos įstatymą“ (Reed, *et al,* 2016, p. 3). „Todėl pasitikėjimo ir atskaitomybės aplinka kuriant ir naudojant dirbtinio intelekto įrenginius ir savarankiškas mokymosi sistemas apima teisinių taisyklių dėl civilinės atsakomybės kūrimą arba esamų taisyklių pritaikymą prie jų naudojimo kylančios rizikos“ (Montagnani, Cavallo, 2021, p. 211).

Naujųjų technologijų ir dirbtinio intelekto atsakomybės klausimus plačiai nagrinėjo Europos Komisijos 2018 metais paskirta Atsakomybės ir naujųjų technologijų ekspertų grupė *(angl. the Expert Grop on Liability and New Technologies, NTF),* kuri 2019 metais pateikė probleminius dirbtinio intelekto ir kitų naujųjų technologijų keliamus atsakomybės aspektus bei pasiūlymus, į ką reikėtų atkreipti dėmesį nagrinėjant būtinus atsakomybės reguliavimo peržiūros elementus.

Šiame skyriuje bus nagrinėjama, kokios mašininio mokymosi algoritmų problemos sąlygoja atsakomybės kvalifikavo klausimus ir kokie aspektai yra svarbūs priimant sprendimus dėl atsakomybės taikymo.

4.1.Atsakomybės kvalifikavimo problematika

Keliant teisines problemas atsakomybės klausimais, kurie yra aktualūs mašininio mokymo sprendimų kontekste reikia pažymėti, kad „atsakomybė grindžiama nuostoliais ar žala, kurią sukėlė koks nors asmuo, veikla ar turtas“ (Reed, *et al*, 2016, p. 4). Toks požiūris vyrauja daugumoje valstybių ir Europos Sąjungos kontekste deliktinė atsakomybė daugeliu atveju nėra harmonizuota išskyrus kelias sritis, t.y. atsakomybė už nekokybiškus produktus pagal Direktyvą 85/374/EC, taip pat jau minėta atsakomybė už duomenų apsaugos apžeidimus pagal BDAR (pagal 82 straipsnį) ir atsakomybė už konkurencijos tesės pažeidimus pagal Direktyvą 2014/104/EU.

Atlikdami ES šalių teisinių sistemų analizę NTF ekspertai padarė išvadą, kad nacionaliniu lygiu šalys neturi atskirų atsakomybės taisyklių, kurios būtų susijusios su dirbtinio intelekto naudojimu. Atskiros šalys naudoja pavienių rinkų reguliavimą, bet daugumoje atvejų bandoma prisitaikyti prie susiformavusių teisinių normų ir rasti tinkamą jų pritaikymą konkrečiam ieškiniui. Pavyzdžiui, „sutarčių lygmeniu informacijos asimetrija, atsirandanti dėl DI naudojimo, gali pateisinti (įstatymų ar teismų praktikos) ikisutartinės atsakomybės režimo taikymą (*culpa in contrahendo* ir panašios sąvokos). Tačiau labiau tikėtina, kad teisinės sistemos reakcija į galimus sutarčių sudarymo, naudojant algoritmus, pažeidimus priklausys nuo sutarčių teisės instrumentų, skirtų sutarčių galiojimui įvertinti ir ginčyti (netinkamas sutikimas, sąžiningumo trūkumas ir kt.)“ (European Commission, 2019, p. 18).

Suprantama, kad dirbtinis intelektas apima daug komponentų susijusių su technologinėmis priemonėmis, pavyzdžiui technologinė infrastruktūra, programinė įranga, duomenys ir duomenų bazės ir kt., o taip pat yra daug susietų suinteresuotų asmenų, kurie dirba kuriant ar naudoja dirbtinio intelekto priemones, pavyzdžiui sprendimo dizaineriai, programuotojai, pardavėjai, naudotojai, vartotojai ir pan. Visi šie elementai tampa svarbūs, kai keliamas klausimas – kas atsakingas už žalą, atsiradusią dėl dirbtinio intelekto sprendimų? „Atsakomybės režimų adekvatumas ir išsamumas technologinių iššūkių akivaizdoje iš tiesų yra labai svarbūs visuomenei. Jei sistema yra netinkama, ydinga arba joje yra trūkumų, susijusių su besivystančių skaitmeninių technologijų padarytos žalos atlyginimu, nukentėjusiesiems gali visiškai arba iš dalies nekompensuota, net jei bendra teisinga analizė gali būti pagrįsta jos atlyginimu“ (Montagnani, Cavallo, 2021, p. 212).

Dauguma autorių, kurie nagrinėja dirbtinio intelekto, mašininio mokymosi sprendimų problematiką, atsakomybės klausimą nagrinėja per deliktinės atsakomybės ir

atsakomybės už produktų kokybę reguliavimą ir elementus, kurie yra būtini tokiai atsakomybei atsirasti. Europos Sąjungos šalyse „vidaus deliktų įstatymai apima taisyklę, nustatančią plačią taikymo sritį kaltės pagrindu, kartu su keliomis konkretesnėmis taisyklėmis, kurios arba pakeičia atsakomybės dėl kaltės prielaidas (ypač paskirstant įrodinėjimo pareigą) arba nustato atsakomybę nepriklausomai nuo kaltės (griežta arba rizika pagrįsta atsakomybė). Dauguma atsakomybės režimų taip pat apima atsakomybės už kitus sąvoką (netiesioginė arba pakaitinė atsakomybė), kuri, savo ruožtu, gali būti (priklausomai nuo atvejo ar šalies) kaltės arba rizikos pagrindu“ (Montagnani, Cavallo, 2021, p. 214). Mašininiu mokymosi pagrįsti sprendimai gali papulti į abi atsakomybės kategorijas, nes pavyzdžiui savaeigio automobilio atveju turime konkretų fizinį produktą, o medicinos diagnostikos atveju mašininis mokymas yra naudojamas sprendimui priimti. Kaip sako Reed „produktai iš esmės yra apčiuopiami arba kilnojamieji objektai, todėl jei mašininio mokymosi technologija įmontuota, tarkime, motorinėje transporto priemonėje, ji tampa to produkto dalimi. Tačiau jei mašininio mokymosi technologija talpinama debesyje, kad jos vartotojai ją gautų kaip paslaugą, atsakomybės už produktą režimas nebus taikomas (Reed, *et al*, 2016, p. 7). Benhamou produkto ar paslaugos klausimą kelia kaip specifinį iššūkį kai kalbama apie dirbtinio intelekto veiksmus ir norime taikyti atsakomybę už produktų kokybę. Autoriaus teigimu, pati produkto sąvoka gali būti interpretuojama labai plačiai, o atsakomybė už produkto kokybę daugiausiai yra siejama su materialiais kilnojamais daiktais, bet ne paslaugomis, o pagrindinės šiuolaikinės technologijos, tokios kaip programinė įranga ir algoritmai yra dažniausiai vertinami kaip paslaugos, o ne produktai. Žalos gali padaryti bet kas, tiek „materialios įrangos (produkto) trūkumai, tiek ji gali kilti dėl nesusipratimų kylančių tarp fizinės infrastruktūros ir dirbtinio intelekto „smegenų“, dėl neteisingos duomenų analizės arba sugadintų trečiosios šalies duomenų, kurie įvedami į DI algoritmą“ (Benhamou, Ferland,, 2020, p. 9), ir doktrininės nuomonės dėl to, ar dirbtinio intelekto programinė įranga turėtų būti kvalifikuojama kaip prekė ar paslauga, nėra. Benhamou sako, kad dirbtinio intelekto sistemos gali būti bet kuo, nuo pasyvių agentų, kurie vykdo specifines žmogaus instrukcijas iki autonominių bendrovių, galinčių mokytis, priimti sprendimus ir veikti nepriklausomai nuo pirminio programavimo, o tai reiškia, kad „toks veikimas gali patenkinti reikalavimus taikomus fiziniam ar juridiniam asmeniui už dirbtinio intelekto veiksmus ir užtikrinti žalos atlyginimą nukentėjusiems asmenims“ (Benhamou, Ferland, 2020, p. 4). Yra specifiniai iššūkiai, kurie esamam atsakomybės reguliavimo režimui yra svarbūs netgi kalbant apie teisinį subjektą, kuriam tektų atsakomybė, tačiau šiai dienai diskusijos, kad dirbtinis intelektas turėtų turėti atskirą teisinio subjektiškumo režimą yra

nevaisingos, nes bet kokiu atveju technologijos prisiriša prie fizinių arba juridinių asmenų.

Kai kalbama apie atsakomybės klausimą dėl žalos ar nuostolių sukeltų mašininio mokymosi, dirbtinio intelekto sprendimais Benhamou išskiria 4 rūšis kliūčių, kurios yra svarbios nagrinėjant atsakomybės atvejus teismuose:

- Didelis kiekis įtrauktų susijusių asmenų;
- Dirbtinio intelekto autonomiškumas (savarankiškumas);
- Paaiškinimo trūkumas;
- Nuspėjamumo trūkumas (Benhamou, Ferland, 2020, p. 5).

Kalbant apie **didelį kiekį įtrauktų susijusių asmenų**, sakoma, kad „skaitmeninės technologijos, tame tarpe ir dirbtinis intelektas tampa labai kompleksiškas dėl tarpusavio priklausomybės tarp jų skirtingų komponentų“ (Benhamou, Ferland, 2020, p. 5), tokių kaip materialios dalys ir įranga, įvairi programinė įranga ir jų aplikacijos, duomenys kaip tokie, duomenų procesai, sujungimo funkcijos. „Kadangi skaitmeninėse ekosistemose dalyvauja daug veikėjų, tampa vis sunkiau išsiaiškinti, kas gali būti atsakingas už padarytą žalą“ (European Commission, 2019, p. 33). Susijusių šalių skaičius, kurie įtraukti į dirbtinio intelekto kūrimą ir veikimą yra nuolat didėjantis: įrangos gamintojai, programinės įrangos dizaineriai, pardavėjai, priemonių ir programinės įrangos diegėjai, priemonių savininkai, dirbtinio intelekto savininkai, dirbtinio intelekto naudotojai ir patikimos trečiosios šalys, tarp kitų, visi gali turėti savo vaidmenį užtikrinant, kad dirbtinis intelektas nepadarytų žalos, todėl priskirti atsakomybę tokiame kontekste nėra lengva užduotis (Benhamou, Ferland, 2020, p. 6). Probleminiai aspektai kyla daugeliu atveju, nes pavyzdžiui atsakomybės už gaminių kokybę atveju galima būtų kalbėti apie solidarią atsakomybę, taip palengvinant atsakomybės klausimą, tačiau dirbtinio intelekto kontekste nuostatos dėl solidariosios atsakomybės gali neaprėpti visų susijusių suinteresuotųjų šalių, net griežtosios atsakomybės atveju būtų sunku nustatyti, kuri dirbtinio intelekto vertės grandinėje esanti šalis turėtų būti laikoma atsakinga, ypač, kai išvados yra priimama dirbtinis intelektas autonomiškai.

Dar vienas papildomas aspektas atsiranda dėl to, kad „skaitmeninės technologijos yra nuolatos modifikuojamos kai jos patenka į rinką per naujų duomenų įtraukimą, programinės įrangos atnaujinimus ir pataisymus, kuriuos taiko dirbtinio intelekto sistemos gamintojas, atskirų sistemos komponentų gamintojai ar net trečiosios šalys“ (Benhamou, Ferland, 2020, p. 6). Tokiais atvejais yra prarandamas ryšys su rizikomis, kurios buvo siejamos su originaliu dirbtinio intelekto produktu, nes nauji kodai ar pakeisti funkcionalumai gali paveikti visos sistemos veikimą, kas gali turėti įtakos viso dirbtinio

intelekto sprendimo saugumui. Kaip rašo Sullyvan, būtų ypatingai nesąžininga griežtosios atsakomybės atveju priskirti kaltę ir jos principus gamintojui ar dizaineriui kurio dirbinys yra labai nutolęs tiek laiko, tiek geografiniu požiūriu nuo originalaus dirbtinio intelekto veikimo (Sullyvan, Schweikart, 2019, p. 163).

Antroji kliūtis dėl atsakomybės priskyrimo yra **dirbtinio intelekto autonomiškumas (savarankiškumas)**. Autoriai sako, kad „šiandieninis dirbtinis intelektas tampa vis savarankiškesnis, nes, nors iš pradžių buvo užprogramuotas žmogaus, dabar jis gali apdoroti duomenis, mokytis iš jų ir priimti nepriklausomus sprendimus, kurie vargu ar gali būti susieti su pradiniu dizainu ar programavimu“ (Benhamou, Ferland, 2020, p. 6). Kalbant apie didelio savarankiškumo dirbtinį intelektą deliktinės atsakomybės atveju reikėtų nustatyti kaip ir kokie subjektai gali pažeisti pareigas arba neteisėtai veikti, kad nustatytume priežastinį ryšį. „Net jei galima manyti, kad asmuo prisidėjo prie žalos padarymo (pvz., programuotojas, kuris nurodė DI atsižvelgti į tam tikros rūšies duomenis), ar jo atsakomybė turėtų būti proporcinga atitinkamo DI savarankiškumo laipsniui ir kaip ar galime tinkamai įvertinti šį savarankiškumo laipsnį? (Giangiacomo, 2018 cituota Benhamou, Ferland, 2020, p. 7). Tokie klausimai keliami žinant, kad dirbtinis intelektas gali veikti ir be žmogaus įsikyšimo, besimokydamas iš savo patirties ir pasikoreguoti, kad pasiektų geresnio efektyvumo. „Jie patys gali pakeisti pradinius algoritmus dėl savarankiško mokymosi galimybių, kurios apdoroja operacijos metu surinktus išorinius duomenis. Tokių duomenų pasirinkimą ir jų įtakos rezultatui laipsnį nuolat koreguoja patys tobulėjantys algoritmai‘ (European Commission, 2019, p. 33).

Kai turime atsakomybės klausimą dėl produktų kokybės, autoriai pažymi, kad esamas reguliavimas nepadengia galimybės teisiškai apibrėžti klaidų, kurias sukelia savarankiškas dirbtinis intelektas. Benhamou pažymi, kad „iš tiesų, daugeliu atveju tiesiog nebus įmanoma nubrėžti linijos tarp žalos, atsiradusios dėl savarankiškų DI sprendimų, ir žalos, atsiradusios dėl gaminio defekto“ (Benhamou, Ferland, 2020, p. 7). „Naujos skaitmeninės technologijos nėra užbaigiamos išleidus į apyvartą, bet pagal savo pobūdį priklauso nuo vėlesnio įnašo, ypač nuo dažnesnių ar retesnių atnaujinimų“ (European Commission, 2019, p. 33). Kai kalbama apie žalos ar nuostolių atlyginimą, ieškovas turėtų įrodyti, kad produktas turėjo trūkumų, tačiau dirbtinio intelekto produktų atveju tai yra sunku įrodyti, nes jie puikiai veikia be jokių mechaninių defektų, tačiau gali sukelti turto sugadinimą ar sužalojimą dėl savo mašininio mokymosi gebėjimų (Benhamou, Ferland, 2020, p. 7). Kaip pabrėžia Benhamou, „kadangi dirbtinis intelektas vystosi ir stiprina save mašininiu mokymusi ir pats prisitaiko, kad taptų „protingesnis“ be

žmogaus įsikišimo, vėliau jis gali tiesiog nebebūti tuo pačiu DI, koks buvo tada, kaip paliko gamintojo rankas, todėl beveik kiekvieną kartą nukentėjusysis paliekamas be kompensacijos“ (Benhamou, Ferland, 2020, p. 7).

Kokį teisių gynimo būdą taikyti žalai ar nuostoliams atlyginti kelia didelių klausimų, nes „tais atvejais, kai dirbtinis intelektas yra skirtas pakeisti žmogaus sprendimų priėmimą, atsakomybės už gaminių kokybės įstatymo normos taikymas „netinkamų“ sprendimų atveju reikštų, kad gamintojai yra atsakingi beveik visais atvejais (dėl griežtos atsakomybės), tačiau žmonėms, darantiems tas pačias klaidas pagal deliktų teisės normas gali būti pripažinta, kad nėra kaltės ar aplaidumo“ (Benhamou, Ferland, 2020, p. 7).

Paaiškinimo trūkumas. Nagrinėjant mašininio mokymosi sampratą aptarta, kad dirbtinio intelekto, mašininio mokymosi veikimas yra paremtas tikslų pasiekimu. „DI dizaineriai neprogramuoja visų galimų scenarijų iš anksto ir neduoda konkrečių nurodymų kiekvienam iš jų; veikiau jie nustato mašinos tikslą ir leidžia dirbtiniam intelektui apdoroti įvestus duomenis, mokytis iš jų ir nuspręsti, kaip geriausiai elgtis, kad būtų pasiektas tikslas“ (Benhamou, Ferland, 2020, p. 8). „Kuo daugiau išorinių duomenų sistemų gali apdoroti ir kuo daugiau jose įdiegta vis sudėtingesnis dirbtinis intelektas, tuo sunkiau numatyti tikslų poveikį, kurį jos turės pradėjus veikti“ (European Commission, 2019, p. 33). Tai reiškia, kad patys DI programuotojai negali paaiškinti proceso kaip „mąsto“ dirbtinis intelektas, kad pasiektų galutinio rezultato, taip pat dirbtinio intelekto klaidos ne visuomet gali būti suprantamos ir paaiškinamos žmonių. Barfield pateikia pavyzdį, kai tiksliosios medicinos algoritmai apdoroja pacientų ir ligoninių duomenis, kad prognozuotų paciento riziką ir suformuluotų diagnozes, tačiau ne visada įmanoma nustatyti, kurie duomenų elementai buvo apdoroti, koks svoris buvo suteiktas kiekvienam elementui atliekant visuotinį vertinimą ir ar yra neetiškas šališkumas apdorojant duomenis (Barfield, 2018, p. 195). „Juodosios dėžės“ dirbtinio intelekto pobūdis sukelia aiškinimo iššūkius ir galiausiai daro įtaką priežastiniam ryšiui ir atsakomybės paskirstymui“ (Benhamou, Ferland, 2020, p. 8). Kalbant apie atsakomybės klausimus „DI sistemos neveikimo priežasties nustatymas yra pagrindinis veiksnys nustatant kaltę ar pareigos elgtis rūpestingai pažeidimą ir priežastinį ryšį deliktiniuose ieškiniuose arba ryšį tarp defekto ir žalos pretenzijose dėl atsakomybės už gaminį“ (Benhamou, Ferland, 2020, p. 8). Esminis dalykas yra tas, kad ieškovas niekada negali sugrįžti į duomenų padorojimo grandinę ir atkurti visą procesą, kad galėtų surasti, kurios aplinkybės lėmė klaidingą išvestį, o tai reiškia, kad įrodymų dėl priežastinio ryšio surinkti beveik neįmanoma.

Paskutinis elementas, sudarantis kliūtis atsakomybės už žalą nustatymui yra **dirbtinio intelekto nuspėjamumo trūkumas**. Dauguma autorių teigia, kad kuo pažangesnis yra dirbtinis intelektas, tuo mažiau prognozuojamas ir nuspėjamas jis tampa, nes funkcionalumas yra vystomas neprižiūrimo mašininio mokymosi pagrindais, kai algoritmai gauna įvesties duomenis be susietų išvesties reikšmių ir yra paliekami savarankiškam mokymuisi iš duomenų, kad pateiktų įdomesnius radinius. Tačiau, kai norima pritaikyti deliktinės atsakomybės teisės principus dirbtinio intelekto veikimui, nenuspėjamumas gali sukelti problemų vertinant kaltės ar rūpestingumo pareigos pažeidimą, o taip pat priežastinį ryšį, nes kai turime nenuspėjamus dirbtinio intelekto veiksmus, tai asmuo kuris juo naudojasi nesitiki, kad bus padaryta žalos kitiems, o tikisi, jog imtasi visų saugumo reikalavimų, kad pažeisimų ar žalos būtų išvengta. „Algoritmai dažnai nebėra kaip daugiau ar mažiau lengvai skaitomas kodas, o kaip „juodoji dėžė“, kuri išsivystė per savarankišką mokymąsi ir kurios poveikį galime išbandyti, bet ne tiek suprasti. Todėl nukentėjusiam darosi vis sunkiau atpažinti tokias technologijas kaip galimą žalos šaltinį, jau nekalbant apie tai, kodėl jos ją sukėlė“ (European Commission, 2019, p. 35). Algoritmo kūrėjai gindami savo poziciją visada patvirtins, kad dirbtinio intelekto sprendimas buvo sukurtas ir ištestuotas, o juo besinaudojantis personalas buvo apmokytas ir prižiūrimas, ir kad visi būtini kokybės kontrolės mechanizmai yra įdiegti (Benhamou, Ferland, 2020, p. 8). „Mašininio mokymosi technologijos, galinčios sukelti fizinius sužalojimus arba žalą turtui, greičiausiai bus pritaikytos tik tada, jei jų bandymai parodys, kad jos užtikrina aukštesnį saugos lygį nei žmogaus sprendimų priėmimas, kurį jos pakeičia. Beveik visada bus tai patvirtinančių įrodymų, ir tokių įrodymų turėtų pakakti įrodyti, kad pagal tą apibrėžimą produktas nėra su trūkumais“ (Reed, *et al*, 2016, p. 6). Tokie patys klausimai kyla ir taikant atsakomybę už produktų kokybę, nors materialaus daikto atveju įrodyti, kad daiktas turėjo defektų atrodytų yra lengviau, tačiau dirbtinio intelekto turinys, esantis materialiame daikte sukelia abejonių, ar būtų galima įrodyti defektą.

Taigi, kaip matome iš aptartų elementų, asmuo, siekdamas įrodyti patirtą žalą turi gana stiprius ribojančius faktorius dėl galimo sėkmingo teisių gynimo, o teismas, vadovaudamasis susiformavusia praktika daugeliu atveju sunkiai galėtų įvertinti visus būtinąsias deliktinės atsakomybės taikymo sąlygas.

4.2.Būtinosios civilinės atsakomybės sąlygos mašininio mokymosi sprendimų kontekste

Mašininio mokymosi sprendimų atveju, norint nustatyti visas būtinas civilinės atsakomybės sąlygas susiduriame su kiekvienos sąlygos ypatumais. Kaip sako Mantagnani „be gerai žinomų atskaitomybės ir skaidrumo trūkumo problemų, naujosios skaitmeninės technologijos meta iššūkį tradicinėms atsakomybės sąvokoms, tokioms kaip žala, priežastinis ryšys ir rūpestingumo pareiga (Montagnani, Cavallo, 2021, p. 217), kurioms skiriamas didelis dėmesys. CK 6.264 straipsnio 1 dalyje numatyta, kad civilinė atsakomybė atsiranda neįvykdžius įstatymuose ar sutartyje nustatytos pareigos (neteisėtas neveikimas) arba atlikus veiksmus, kuriuos įstatymai ar sutartis draudžia atlikti (neteisėtas veikimas), arba pažeidus bendro pobūdžio pareigą elgtis atidžiai ir rūpestingai.

Pagal skirtingas civilinės atsakomybės teisės tradicijas gali skirtis kaip aiškinami neteisėti veiksmai ir kaltė, ar jos aiškinamos kaip viena sąlyga (Anglija), ar kaip dvi savarankiškos (Lietuva). „Nesvarbu, ar teisės sistemoje išskiriami objektyvūs ar subjektyvūs neteisėti veiksmai ir (arba) atsakomybės už nusižengimą pagrindas skirstomas į neteisėtumą ir kaltę, išlieka du esminiai dalykai: nustatyti rūpestingumo pareigas, kurias kaltininkas turėjo atlikti, ir įrodyti, kad kaltininko elgesys yra netinkamas ir kaltininkas šių pareigų nevykdė“ (European Commission, 2019, p. 23). Šiuos du elementus įrodyti gali būti sunku, nes „fiziniams ir juridiniams asmenims, patyrusiems žalą dėl naujųjų skaitmeninių technologijų produktų, gali trūkti priemonių patikrinti galimus įstatymų pažeidimus, o tai trukdo veiksmingai kreiptis į teismą“ (Montagnani, Cavallo, 2021, p. 216). Kalbant apie mašininį mokymąsi dažniausiai susiduriama su paslaugomis, nes „produktai yra fiksuoto pobūdžio ir, jei technologijos mašininio mokymosi elementas reguliariai keičiasi naudojant išorinius naujinimus, tikėtina, kad jie bus parduodami kaip paslauga ir pagal įstatymą traktuojami kaip tokie“ (Reed, *et al*, 2016, p. 7). Daugeliu atveju susidursime su pareigos elgtis atidžiai ir rūpestingai pažeidimu. „Dirbtinio intelekto sistemose vykdomų procesų negalima išmatuoti atsižvelgiant į rūpestingumo pareigas, skirtas žmogaus elgesiui, arba be koregavimų, kuriuos reikėtų papildomai pagrįsti“ (European Commission, 2019, p. 23).

Rūpestingumo pareiga taikoma žmogui, todėl čia negalime taikyti *bonus pater familias* kriterijaus, kad įvertintume mašininio mokymosi sprendimo standartą. Kalbant apie technologijas teisėje „keliami ir kiti klausimai, pavyzdžiui, ar technologijos gamintojas pakankamai rūpinosi ją projektuodamas, konstruodamas ar išbandydamas, ar technologijos naudotojo sprendimas priimti ją buvo pagrįstas ir ar ji buvo eksploatuojama

tinkamai priežiūra ir įgūdžiai. Kitaip tariant, ji ieško žmogiškųjų, o ne mašinų nesėkmių“ (Reed, *et al,* 2016, p. 10). Kadangi Europos Sąjungos teisė daugiausia reguliuoja produktų ir saugumo reikalavimus, todėl tikėtina, kad ateityje bus įvedamos taisyklės, kurios padėtų apibrėžti rūpestingumo pareigas, tokias kaip prisijungimo registravimo reikalavimai (*angl. logging requirements*), kad peržiūrėti kas galėjo būti padaryta, susijusias su deliktine atsakomybe, jei atrastų žala. „Tokių įstatymų ar teisės aktų reikalavimų pažeidimas taip pat gali lengviau sukelti nukentėjusiojo atsakomybę, pavyzdžiui, daugelyje sistemų perkeliant kaltės įrodinėjimo naštą. Visgi tokių reikalavimų nebus nuo pat pradžių ir gali prireikti metų, kol tokios taisyklės pasirodys teisės aktuose ar teismuose“ (European Commission, 2019, p. 23).

Kaltės nustatymo atveju CK 6.248 straipsnis sako, kad asmuo kaltas, jei atsižvelgiant į prievolės esmę ir kitas aplinkybes jis nebuvo tiek rūpestingas ir apdairus, kiek atitinkamomis sąlygomis buvo būtina. Kaltė gali pasireikšti tyčia arba neatsargumu, o tam tikrais atvejais atsakomybė atsiranda be kaltės. Mašininio mokymosi ir dirbtinio intelekto technologijos padarytos žalos atveju, gali kilti problemų įrodant kaltę. „Paprastai nukentėjusysis turi įrodyti, kad kaltinamasis (arba asmuo, kurio elgesys jam priskirtinas) buvo kaltas. Todėl nukentėjusysis turi ne tik nustatyti, kokias rūpestingumo pareigas kaltinamasis turėjo atlikti, bet ir įrodyti teismui, kad šios pareigos buvo pažeistos. Įrodžius atsakovo kaltę, reikia pateikti teismui įrodymus, kurie leistų jam manyti, koks buvo taikomas rūpestingumo standartas ir kad jo nesilaikoma“ (European Commission, 2019, p. 24), taip pat reikia „pateikti įrodymus, kaip įvyko įvykis, dėl kurio buvo padaryta žala. Kuo sudėtingesnės aplinkybės, dėl kurių aukai buvo padaryta žala, tuo sunkiau nustatyti atitinkamus įrodymus“ (European Commission, 2019, p. 24). Toks procesas gali būti labai sudėtingas ir brangus, pavyzdžiui nustatant klaidą ilgame programiniame kode, arba išnagrinėti procesą vedantį prie konkretaus rezultato, kaip įvesties duomenys atvedė prie išvesties rezultato. Kadangi algoritmai pasižymi mokymosi savybe, nustatyti, ar tikrai tai buvo būtent algoritmo kaltė, ar kalti veiksmai privedę prie žalos, sukeltos algoritmais. „Kalbant apie besivystančias skaitmenines technologijas, nusistovėjusių šių technologijų tinkamo veikimo modelių trūkumas ir tai, kad jos vystosi mokantis be tiesioginės žmogaus kontrolės, apsunkina kaltės pagrįstos atsakomybės taisyklių taikymą. Nors dirbtinio intelekto sistemose veikiantys procesai negali būti išmatuoti pagal žmogaus elgesio pareigas, priimtas autonominių sistemų kūrimo ir eksploatavimo priežiūros standartas dar neatsirado“ (Montagnani, Cavallo, 2021, p. 218).

Deliktinės atsakomybės tikslas yra atlyginti nukentėjusiam asmeniui nuostolius ar žalą, kuriuos jie patyrė ir neturėtų patys prisiimti. Tačiau žala atlyginama tik tiems suinteresuotiems asmenims, kurie patenka po teisinės sistemos apsauga.

Doktrinoje „vieningai sutariama, kad asmens ar fizinio turto sužalojimai gali sukelti deliktinę atsakomybę“ (European Commission, 2019, p. 19), tačiau nėra vieningos pozicijos dėl grynai ekonominių nuostolių. „Pavyzdžiui, finansų rinkose besimokančių algoritmų padaryta žala dažnai liks neatlyginta, nes kai kurios teisinės sistemos neužtikrina tokių interesų deliktinės apsaugos iš viso arba tik tuo atveju, jei įvykdomi papildomi reikalavimai, pavyzdžiui, sutartiniai santykiai tarp šalių arba pažeidžiamos kokios nors konkrečios elgesio taisyklės“ (European Commission, 2019, p. 19). „Kalbant apie žalos sąvoką, be tradicinės žalos (žalos asmenims ir turtui), yra ir tokių, susijusių su duomenų perdavimu, privatumu ir konfidencialios informacijos saugumu“ (Montagnani, Cavallo, 2021, p. 217). Tačiau dėl pačios žalos kompensuotinumo klausimo doktrinoje nėra, labiau keliamas klausimas dėl žalos kategorijų, kurios gali nepapulti po tradicinėmis patirtos žalos kategorijomis, dėl naujų technologijų poveikio. „Dabar, nors asmens ar fizinio turto sužalojimai gali sukelti atsakomybę, vien ekonominių nuostolių atlyginimas nėra visuotinai priimtas, o duomenų sunaikinimas taip pat nėra turto praradimas. Taip pat scenarijuje, kai neigiamai paveikiamos asmens teisės, pavyzdžiui, kai duomenys išplatinami pažeidžiant teisę į privatumą, jurisdikcijose yra skirtumų“ (Montagnani, Cavallo, 2021, p. 217). Kadangi, „nei viena valstybė narė nėra apibrėžusi nuosavybės teisės į duomenis sąvokos“ (Medvedevaitė, Mickevičiūtė, 2021) ir savarankiškai reguliuoja duomenų valdymo ir teisinio statuso klausimus, tai įvertinti atsiradusio žalos dydį ir pasekmes tampa sudėtinga.

„Tačiau labiausiai ginčytinas atsakomybės režimo elementas yra priežastinis ryšys tarp nukentėjusiajam padarytos žalos ir kaltinamojo veiksmų. Iš esmės pagal deliktinę teisę nukentėjusysis turėtų įrodyti, kad žala atsirado dėl tam tikro atsakovo elgesio ar rizikos“ (Montagnani, Cavallo, 2021, p. 217). CK 6.247 straipsnis sako, kad atlyginami tik tie nuostoliai, kurie susiję su veiksmais (veikimu, neveikimu), nulėmusiais skolininko civilinę atsakomybę tokiu būdu, kad nuostoliai pagal jų ir civilinės atsakomybės prigimtį gali būti laikomi skolininko veiksmų (veikimo, neveikimo) rezultatu. Ieškovo galimybės identifikuoti santykį tarp žalos ir neteisėto veikimo mašininio mokymosi sprendimo atveju pareikalautų itin stiprios kvalifikacijos, nes mašininio mokymosi pradinės informacijos beveik neįmanoma atsekti. „Pateikti priežastinio ryšio įrodymų yra dar sunkiau, kai susiduriama su savarankiškai besimokančiomis DI sistemomis, kurias skatina mašininis mokymasis ir gilaus mokymosi metodai ir, kurie remiasi daugybe išorinių

duomenų surinkimu“ (Montagnani, Cavallo, 2021, p. 217). „Net ir nepakeitus pradinės programinės įrangos dizaino, įterptieji kriterijai, kuriais vadovaujamasi renkant ir analizuojant duomenis bei sprendimų priėmimo procesą, gali būti sunkiai paaiškinami ir dažnai reikalauja brangios ekspertų analizės“ (European Commission, 2019, p. 20). Todėl šiandieninė problema susijusi su priežastinio ryšio nustatymu dėl didelio sistemų veikimo neapibrėžtumo ir įrodinėjimo standartų trūkumu. Jeigu atsirastų nusistovėję minimalūs reikalavimai sistemų veikimui, jais teismai galėtų pasiremti sprendžiant klausimus dėl priežastinio ryšio nustatymo ir taip suteikiant galimybę pateikti įrodymus nukentėjusiajam. Visgi dabar teismai turės taikyti arba įstatymo analogiją, arba nustatyti minimalius įrodinėjimo reikalavimus, paremtus privalomaisiais nurodymais ir taisyklėmis. „Griežtos atsakomybės režimu toks įrodinėjimas galėtų būti mažiau problemiškas, nes jo pakaktų įrodyti, kad rizika, dėl kurios atsiranda griežtoji atsakomybė, išsipildė, tačiau griežta atsakomybė taikoma tik labai ribotais atvejais“ (Montagnani, Cavallo, 2021, p. 217).

Apibendrinant civilinės atsakomybės sąlygų nustatymo problematiką pagrindiniai kritiniai aspektai yra susiję su neapibrėžtumu, kurį sukelia naudojamos technologijos ir visuotinai priimtinių elgesio standartų nebuvimu. Neabejotinai, esamos teisės normos ilguoju laikotarpiu turės būti modifikuojamos, siekiant atliepti technologijomis grindžiamų civilinių santykių pokyčius. Visgi jau dabar teisės srities mokslininkai (Benhamou, European Commission) skirtingose jurisdikcijose siūlo peržvelgi esamus atsakomybės taikymo režimus ir nekuriant naujų atsakomybės principų pritaikyti esamas sąlygas mašininio mokymosi, dirbtinio intelekto sprendimams per padidintą reikalavimą rūpestingumo pareigoms ir tikslumui, daline ir solidariąją atsakomybe už žalą. Taip pat peržiūrint galimybes palengvinti nukentėjusiesiems įrodymų surinkimo naštą, nustatant minimalius reikalavimus, kuriais galima būtų remtis įrodant būtinųjų pareigų nesilaikymą (pavyzdžiui, bendrovė turi pateikti privalomus įrašus ar informaciją ir jeigu nepateikia, laikoma, kad ji jų neturi). Taip pat ir žalos nustatymo ir apskaičiavimo klausimai gali būti įvertinami per esamus reguliacinius modelius taikomus atskiroms sritims (vartotojų apsauga, asmens duomenų apsauga, privatumo apsauga, intelektinės nuosavybės apsauga, konkurencijos apsauga, sutarčių teisės normos ir pan.). Visą tai galima būtų potencialiai įgyvendinti per besiformuojančią teismų praktiką skirtingose jurisdikcijose.

IŠVADOS

1. Mašininio mokymosi sprendimų problematiką apsprendžia du svarbiausi elementai tai – mokymosi duomenys ir algoritmai. Pagal savo veikimo principą mašininis mokymasis yra viena iš dirbtinio intelekto posistemių, kai algoritmai mokomi nustatyti tam tikrus dėsningumus iš jiems duoto duomenų rinkinio, pagal kuriuos nustatoma, kokių veiksmų reikia konkrečiam tikslui pasiekti. Naudodami didelius duomenis mašininio mokymosi algoritmai didina savo našumą, taip pakeisdami žmogaus veikimą. Veikdami neprižiūrimoje aplinkoje algoritmai tampa savarankiški ir tampa neįmanoma nustatyti, kuriuos įvesties duomenis naudojo mašininio mokymosi modelis, kad pateiktų išvestį (rezultatą). Dėl šios priežasties atsiranda „juodosios dėžės“ problematika, kuri sukelia teisinius ir etinius iššūkius, nes technologijoms įsivyraujant privačiuose santykiuose sprendimų priėmimas tenka savarankiškam ir sunkiai kontroliuojam objektui.

2. Didžiausias dėmesys mašininio mokymosi algoritmų naudojime turi būti skiriamas mokymosi duomenų kokybei. Šiandien duomenų industrija leidžia sekti įvairiausius duomenis per išmaniuosius įrenginius, operacijas ir taip juos susieti su asmeniu, todėl atsiranda galimybės naudoti šališkus duomenis, kurie privataus asmens atžvilgius kelia sąžiningumo ir diskriminavimo problemas. Įmonės naudodamos profiliuotus duomenis gali pasinaudoti vartotojų polinkiais ir siūlyti prekes ir paslaugas atitinkančias jų pageidavimus ir už tokią kainą, kurią jie nori mokėti, tačiau asmuo gali nežinoti apie individualiai pritaikomą kainodarą arba tikslinę rinkodarą, taip pažeidžiant sąžiningumo principą. Taip pat naudojami praeities mokomieji duomenys gali turėti šališkumą ir diskriminuoti vartotojus lyties, rasės, pajamų ir kitais aspektais, taip pažeidžiant lygiateisiškumo principą.

3. Mašininio mokymosi algoritmo skaidrumas yra esminė problema kylanti iš paaiškinamumo trūkumo, priėjimo prie informacijos disbalanso. Paaiškinti algoritmo veikimą gali būti sunku arba neįmanoma, dėl nuolatinio algoritmo mokymosi proceso, o žmogus ne visuomet gali suprasti, kokį sprendimą priima algoritmas. O jeigu ir supranta, ne visuomet gali su sprendimu sutikti. Dėl sudėtingos algoritmų veikimo logikos, algoritmų kūrėjai, savininkai gali pasinaudoti prieigos prie informacijos galimybėmis ir siekti savo komercinių interesų, taip įgydami pranašumą prieš vartotoją. Todėl siekiant apsaugoti asmenų interesus, būtina stiprinti algoritmų atskaitomybės reikalavimus, kaip galimybę pasiekti interesų pusiausvyros tarp algoritmų savininkų ir vartotojų. Vienas iš būdų – nustatyti bent minimalius standartus ir taisykles, kurių privaloma laikytis ir kurių nesilaikymas leistų preziumuoti nesąžiningą elgesį.

4. Naudojamų duomenų kiekis mašininio mokymosi sprendimuose apima įvairiausio pobūdžio duomenis apie asmenį, kurių pagalba stebimi ir profiliuojami asmenys. Dėl tokio srauto informacijos, kyla asmens duomenų apibrėžties klausimas, nes BDAR saugo asmens duomenis, tačiau neapima iš tų duomenų padarytų išvadų ir elgesio prognozių. Naujų duomenų generavimas ir naudojimas šiandien nėra aiškiai reglamentuotas, nes BDAR saugo duomenų įvesties pusę, o išvestiniai duomenys, iš kurių daromos įžvalgos ir prognozės yra „pilkoji zona“ ir apsaugoti silpnai. Doktrinoje pažymima, kad išvestiniai duomenys turėtų būti asmens duomenimis ir patekti po asmens duomenų apsauga. Tuo tarpu Europos Teisingumo Teismas vienareikšmiškos nuomonės nepateikia ir rekomenduoja spręsti apie išvestinių duomenų priskirimą asmens duomenims pagal kontekstą. Tokia teismo pozicija atveria erdvę interpretacijoms ir sukelia teisinį netikrumą, dėl galimybės pasinaudoti išimtimis ir kitų institutų teikiama apsauga, pavyzdžiui intelektinės nuosavybės ar komercinių paslapčių įstatymais.

5. Siekiant didinti vartotojų pasitikėjimą skaitmeninėmis paslaugomis ir apsaugoti vartotojus nuo pažeidimų, naujas vartotojų teisių apsaugos reguliavimas suteikia naujas galimybes kontroliuoti profiliais ir išvestiniais duomenimis paremtą turinį vartotojams. Ypatingą reikšmę Omnibus direktyvos įgyvendinimui turės galimybė atsiskaityti už paslaugas ne pinigais, o asmens duomenimis. Asmens duomenys įgyja prielaidas būti turtinių santykių objektu ir dalyvauti civilinėje apyvartoje. Lietuvoje asmens duomenys pagal CK nėra laikomi civilinės teisės objektais, tačiau neišvengiamai turi įvykti pokyčiai dėl asmens duomenų traktavimo, nes tai keičia ir atlygintinumo principo traktavimą sutarčių teisėje.

6. Mašininiu mokymusi priimami sprendimai keičia civilinės atsakomybės sąlygų įrodinėjimo standartus. Kompleksiškumas, tarpusavio priklausomybė ir didelis įtrauktų asmenų kiekis į mašininio mokymosi algoritmų kūrimą ir valdymą sąlygoja kaltų asmenų nustatymo kliūtis. Dėl nematerialaus mašininio mokymosi žalos pobūdžio yra sudėtinga įrodyti ir apskaičiuoti žalą. Mašininio mokymosi algoritmams taikomų standartų trūkumas sąlygoja rūpestingumo pareigos pažeidimą, nes rūpestingumo pareiga taikoma žmogui, o ne algoritmui. Algoritmų neskaidrumas labiausiai daro įtaką priežastinio ryšio tarp neteisėto elgesio ir patirtos žalos nustatymui. Dėl mašininio mokymosi technologijos nuolatinės evoliucijos šiandieninis civilinės atsakomybės reguliavimas turi būti pritaikomas atsižvelgiant į naujų skaitmeninių technologijų keliamus iššūkius, tačiau nusistovėjęs reguliavimas galėtų būti keičiamas ilguoju laikotarpiu, o praktiniai aspektai turėtų būti įvertinti per naujai besiformuojančią tesimų praktiką.

ŠALTINIŲ SĄRAŠAS

Teisės norminiai aktai

Europos Sąjungos reglamentai ir direktyvos:

1. Europos Parlamento ir Tarybos reglamentas (ES) 2016/679. 2016 m. balandžio 27 d. dėl fizinių asmenų apsaugos tvarkant asmens duomenis ir dėl laisvo tokių duomenų judėjimo ir kuriuo panaikinama Direktyva 95/46/EB (Bendrasis duomenų apsaugos reglamentas). OL L 119, 2016 5 4, p. 1—88. http://data.europa.eu/eli/reg/2016/679/oj

2. Europos Parlamento ir Tarybos direktyva (ES) 2019/2161 2019 m. lapkričio 27 d. kuria iš dalies keičiamos Tarybos direktyva 93/13/EEB ir Europos Parlamento ir Tarybos direktyvos 98/6/EB, 2005/29/EB ir 2011/83/ES, kiek tai susiję su geresniu Sąjungos vartotojų apsaugos taisyklių vykdymo užtikrinimu ir modernizavimu. OL L 328, 2019 12 18, p. 7—28. http://data.europa.eu/eli/dir/2019/2161/oj

3. Tarybos Direktyva 1985 m. liepos 25 d. dėl valstybių narių įstatymų ir kitų teisės aktų, reglamentuojančių atsakomybę už gaminius su trūkumais, derinimo (85/374/EEB). OL L 210, 1985 8 7, p. 29—33. http://data.europa.eu/eli/dir/1985/374/oj

Lietuvos Respublikos teisės aktai:

4. Lietuvos Respublikos civilinis kodeksas. Valstybės žinios, 2000, nr. 74 – 2262.

Specialioji literatūra:

Knygos

5. O'NEIL, Cathy. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Elektroninės knygos:

6. WACHTER S., MITTELSTADT B. (2019). *A right to reasonable inferences: re-thinking data protection law in the age of Big data and* AI. [interaktyvus] Columbia Business Law Review – Vol.– Issue 2. [interaktyvus] Prieiga per internetą: https://ssrn.com/abstract=3248829 [žiūrėta 2022 m. kovo 12 d.]

Straipsniai elektroniniuose mokslo žurnaluose:

7. BARFIELD Woodrow. (2018) *Liability for Autonomous and Artificially Intelligent Robots* [interaktyvus] Paladyn Journal of Behavioral Robotics. Volume 9, Issue 1, pp. 193-203; Prieiga per internetą:

https://www.degruyter.com/document/doi/10.1515/pjbr-2018-0018/html [žiūrėta 2022 m. kovo 10 d.]

8. BENHAMOU Yaniv and FERLAND Justine, (2020) *Artificial intelligence & damages: assessing liability and calculating damages*. [interaktyvus] Prieiga per internetą: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3535387 [žiūrėta 2022 m. kovo 1 d.]

9. BURRELL Jenna, (2016). *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*. [interaktyvus] Sage journals. Big Data & Society 3, no.1, January 5: 1–12. Prieiga per internetą: https://doi.org/10.1177/2053951715622512 [žiūrėta 2022 m. kovo 12 d.]

10. GALLI Federico, (2020). *AI and consumer manipulation: what is the role of EU fair marketing law?* [interaktyvus] Catolica law review. Volume IV, N.2., 35-64. Prieiga per internetą: https://doi.org/10.34632/catolicalawreview.2020.9320 [žiūrėta 2022 m. kovo 19 d.]

11. GROCHOWSKI M., JABŁONOWSKA A., LAGIOIA F. & SARTOR G, (2021). *Algorithmic Transparency and Explainability for EU Consumer Protection: Unwrapping the Regulatory Premises*. [interaktyvus] Critical Analysis of Law (CAL) Vol. 8, No 1, 43-63. Max Planck Private Law research Paper No. 21/7. Prieiga per internetą: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3826415 [žiūrėta 2022 m. kovo 19 d.]

12. MACMILLAN R. (2019). *Big Data, Machine Learning, Consumer Protection and Privacy*. [interaktuvus] TPRC47: The 47th Research Conference on Communication, Information and Internet Policy. July 26. Available at SSRN. Prieiga per internetą: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3427206 [žiūrėta 2022 m. vasario 24 d.]

13. MEDVEDEVAITĖ A., MICKEVIČIŪTĖ G.V. (2021). Skaitmeniniai duomenys: ar reikalingas naujas teisinis reguliavimas. [interaktyvus] Prieiga per internetą: https://doi.org/10.15388/TMP.2021.2 [žiūrėta 2022 m. kovo 19 d.]

14. MONTAGNANI, Maria L. and CAVALLO, Mirta. (2021). "*Liability and Emerging Digital Technologies: An EU Perspective*" [interaktyvus] Notre Dame Journal of International & Comparative Law: Vol. 11 : Iss. 2 , Article 4. Prieiga per internetą: https://scholarship.law.nd.edu/ndjicl/vol11/iss2/4 [žiūrėta 2022 m. kovo 12 d.]

15. REED Chris, KENNEDY Elizabeth, NOGUEIRA Silva Sara, (2016). *Responsibility, Autonomy and Accountability: Legal Liability for Machine Learning*. [interaktyvus] Queen Mary University of London, School of law Legal Studies Research

Paper No. 243/2016 Prieiga per internetą: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2853462 [žiūrėta 2022 m. kovo 1 d.]

16. RODRIGUES Rowena, (2020). *Legal and human rights issues of AI: Gaps, challenges and vulnerabilities*. [interaktyvus] Journal of Responsible Technology 4 (2020) 100005. Prieiga per internetą: www.elsevier.com/locate/jrt [žiūrėta 2022 m. vasario 16 d.]

17. SO, Anthony. (March 30, 2020). *Technical Elements of Machine Learning for Intellectual Property Law*. [interaktyvus] Artificial Intelligence and Intellectual Property. SSRN Prieiga per internetą:

18. https://ssrn.com/abstract=3635942 or http://dx.doi.org/10.2139/ssrn.3635942 [žiūrėta 2022 m. kovo 1 d.]

19. SULLIVAN and SCHWEIKART Scott J.. (February, 2019). "*Are Current Tort Liability Doctrines Adequate for Addressing Injury Caused by AI?*" [interaktyvus] AMA Journal of Ethics, Vol. 21, no. 2, pp. 160-166. https://journalofethics.ama-assn.org/article/are-current-tort-liability-doctrines-adequate-addressing-injury-caused-ai/2019-02 [žiūrėta 2022 m. kovo 19 d.]

20. SURDEN Harry. (2014). *Machine Learning and Law*. [interaktyvus] 89 WASH. L. REV. 87. Prieiga per internetą: https://scholar.law.colorado.edu/articles/81. [žiūrėta 2022 m. kovo 20 d.]

21. TENE, Omer; POLONETSKY, Jules. (2012). *To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioral Advertising*. [interaktyvus] University of Minnesota. Consortium on Law and Values in Health, Environment & the Life Sciences. Retrieved from the University of Minnesota Digital Conservancy. Prieiga per internetą: https://hdl.handle.net/11299/155947 [žiūrėta 2022 m. kovo 28 d.]

22. TURING A.M. (1950). *Computing machinery and intelligence*. [interaktyvus]. Mind, Volume LIX, Issue 236, Pages 433–460. Prieiga per internetą: https://doi.org/10.1093/mind/LIX.236.433 [žiūrėta 2022 m. kovo 1 d.]

23. WILLIAMSON, Ben. 2016. *Computing Brains: Learning Algorithms and Neurocomputation in the Smart City*. [interaktyvus] Information, Communication & Society 0(0):1–19. Prieiga per internetą:

24. https://www.researchgate.net/publication/301775259_Computing_brains_learning_algorithms_and_neurocomputation_in_the_smart_city [žiūrėta 2022 m. kovo 19 d.]

Disertacija:

25. SIAPKA Anastasia, (2019). *The Ethical and Legal Challanges of Artificial Intelligence: the EU response to biased and discriminatory AI*. [interaktyvus] Disertacija. Panteion University of Athens. Parengata: gruodžio 11, 2018, publikuota birželio 27, 2019. https://papers.ssrn.com/sol3/ppers.cfm?abstract_id=3408773 [žiūrėta 2022 m. kovo 1 d.]

Teismų praktika

Europos Sąjungos teisingumo teismo sprendimai:

26. Peter Nowak v. Data Protection Commissioner [ESTT], Nr. C-434/16, [2017-12-20], ECLI:EU:C:2017:994

27. YS, M and S v. Minister voor Immigratie, Integratie en Asiel [ESTT], Jungtinė byla Nr. C-141 ir 372/12. [2014-07-14], ECLI:EU:C:2014:2081

Kiti šaltiniai

***Travaux préparatoires*:**

28. 2018 m. balandžio 11 d. Europos Komisijos komunikatas Europos Parlamentui, Tarybai, Europos ekonomikos ir socialinių reikalų komitetui ir regionų komitetui. Naujos galimybės vartotojams. Briuselis. COM(2018) 183 final. Prieiga per internetą: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1573718927782&uri=CELEX%3A52018DC0183

29. 2018 m. balandžio 11 dienos Europos komisijos pasiūlymas dėl Europos Parlamento ir Tarybos direktyvos kuria iš dalies keičiamos 1993 m. balandžio 5 d. Tarybos direktyvos 93/13/EEB, Europos Parlamento ir Tarybos direktyvos 98/6/EB, Europos Parlamento ir Tarybos direktyvos 2005/29/EB ir Europos Parlamento ir Tarybos direktyvos 2011/83/ES nuostatos, susijusios su geresniu ES vartotojų apsaugos taisyklių vykdymo užtikrinimu ir modernizavimu. {SWD(2018) 96} - {SWD(2018) 98}. Briuselis. COM(2018) 185 final 2018/0090 (COD) Prieiga per internetą: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018PC0185&from=GA

30. 2019 m balandžio 8 dienos Europos Komisijos komunikatas Europos Parlamentui, Tarybai, Europos ekonomikos ir socialinių reikalų komitetui ir regionų komitetui. Pasitikėjimo į žmogų orientuotu dirbtiniu intelektu didinimas. Briuselis. COM(2019) 168 final. Prieiga per internetą: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2019:168:FIN

31. 2020 m. gruodžio 15 d. Europos Komisijos pasiūlymas dėl Europos Parlamento ir Tarybos reglamento dėl bendrosios skaitmeninių paslaugų rinkos (Skaitmeninių paslaugų aktas), kuriuo iš dalies keičiama Direktyva 2000/31/EB. Briuselis. COM(2020) 852 final 2020/0364(COD). Prieiga per internetą: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32020R0852

32. 2020 m. vasario 19 d. Europos Komisijos Baltoji knyga. Dirbtinis intelektas. Europos požiūris į kompetenciją ir pasitikėjimą. Briuselis, 2020 02 19 COM(2020) 65 final [interaktyvus]. [žiūrėta 2022 m. vasario 28 d.]. Prieiga per internetą: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligencefeb2020_en.pdf

33. 2021 m. balandžio 21 dienos Europos Komisijos pasiūlymas dėl Europos Parlamento ir tarybos Reglamento kuriuo nustatomos suderintos dirbtinio intelekto taisyklės (Dirbtinio intelekto aktas) ir iš dalies kaičiami tam tikri Sąjungos teisėkūros procedūra priimti aktai. {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}. Briuselis. COM(2021) 206 final 2021/0106 (COD). Prieiga per internetą: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:52021PC0206

Europos Sąjungos it tarptautinių institucijų leidiniai:

34. ALGORITHMS AND HUMAN RIGHTS. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications. Prepared by the committee of experts on internet intermediaries (MSI-NET) DGI(2017)12, Council of Europe, March 2018. Prieiga per internetą: https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5

35. Boucher Phillip Nucholas. (June, 2020). *Artificial intelligence: How does it work, why does it matter, and what can we do about it?* [interaktyvus] European Parliament. European Parliamentary Research Service. Scientific Foresight Unit (STOA) PE 641.547 Prieiga per internetą: https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641547

36. European Commission, Directorate-General for Justice and Consumers (2019). *Liability for artificial intelligence and other emerging digital technologies*, Publications Office. Prieiga per internetą: https://data.europa.eu/doi/10.2838/25362

37. Europos Komisija, Ryšių tinklų, turinio ir technologijų generalinis direktoratas, *Patikimo DI etikos gairės*, Leidinių biuras, 2019. Prieiga per internetą: https://data.europa.eu/doi/10.2759/55717

38. OECD (2021). *Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers*. [ineteraktyvus] Prieiga per internetą: https://www.oecd.org/finance/artificial-intelligence-machine-learningbig-data-in-finance.htm. [žiūrėta 2022 m. kovo 19 d.]

39. OECD (2018). *The regulation of personalized pricing in the Digital era* – Note by Marc Bourreau and Alexandre de Strell. [interaktyvus] DAF/COMP/WD (2018)150. Prieiga per internetą: https://ssrn.com/abstract=3312158 [žiūrėta 2022 m. kovo 19 d.]

40. REILLON Vincent, (2018). *Understanding Artificial Intelligence*. Briefing. [interaktyvus] European Parliamentary Research Service, European Parliament, January. Prieiga per internetą: http://www.iberglobal.com/files/2018/Understanding_AI.pdf [žiūrėta 2022 m. kovo 19 d.]

41. SARTOR Giovanni, LAGIOIA Francesca, and DG, EPRS. (June, 2020). The impact of General Data Protection Rugulation (GDPR) on artificial intelligence. [interaktyvus] European Parliament. European Parliamentary Research Service. Scientific Foresight Unit (STOA) PE 641.530 https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641530 [žiūrėta 2022 m. kovo 12 d.]

Informaciniai leidiniai, interneto portalai

42. BINNS Reuben, LYNGS Ulrik, VAN KLEEK Max, ZHAO Jun, LIBERT Timothy, SHADBOLT Nigel. (18 October, 2018). *Third Party Tracking in the Mobile Ecosystem*. [interaktyvus] Xiv:1804.03603v3 [cs.CY]. Prieiga per internetą: https://arxiv.org/pdf/1804.03603.pdf; [žiūrėta 2022 m. kovo 1 d.]

43. GIANGIACOMO Olivi, MIELE Claudio Orlando and SCHIAVO Valeria, (2018) *Robots and Liability: who is to blame?* [interaktyvus] Dentons, December 20. Prieiga per internetą: https://www.dentons.com/en/insights/articles/2018/december/20/robots-and-liability [žiūrėta 2022 m. kovo 19 d.]

44. GRIFFIN, Andrew, (2016). *How Facebook Is Manipulating You to Vote*. [interaktyvus] The Independent. Prieiga per internetą: http://www.independent.co.uk/life-style/gadgets-and-tech/news/uk-elections-2016-how-facebook-is-manipulating-you-to-vote-a7015196.html [žiūrėta 2022 m. vasario 22 d.]

45. LAZAUSKAITĖ R. (2021.09.20). *Naujoji „Omnibus“ vartotojų apsaugos direktyva: ką turėtų žinoti kiekvienas elektroninėje erdvėje prekiaujantis verslininkas?*

[interaktyvus] Prieiga per internetą: https://www.glimstedt.lt/naujienos/naujoji-omnibus-direktyva/ [žiūrėta 2022 m. kovo 19 d.]

46. RAM Aliya, WISNIEWSKA Aleksandra, KAO Joanna S., RININSLAND Ændrew, NEVITT Caroline. (2018). *How smartphone apps track users and share data*. [interaktyvus] Financial Times, 23 October. Prieiga per internetą: https://ig.ft.com/mobile-app-data-trackers/ [žiūrėta 2022 m. kovo 12 d.]

47. SKIAUTERIENĖ, D. (2020). *Kaip apsaugoti savo teisę į privatumą internete?* [interaktyvus] Žurnalas Teismai.lt Nr 3(39), Spalis, 2020. ISSN 2029-9451. Prieiga per internetą: https://www.teismai.lt/lt/naujienos/zurnalas-teismai.lt/124/2020m [žiūrėta 2022 m. kovo 19 d.]

SANTRAUKA

Mašininiu mokymu grįstų sprendimų keliami svarbiausi dabartiniai teisiniai iššūkiai

Gytautė Peseckaitė-Kibickienė

Magistro darbe analizuojama teisinė problematika susijusi su mašininio mokymosi sprendimais, kuriuos išskiria vyraujantis politinis diskursas ir teisės mokslo doktrina. Svarbiausi dabartiniai teisiniai iššūkiai nagrinėjami per asmens privatumo apsaugos, vartotojų apsaugos ir atsakomybės už žalą elementus, kurie pasireiškia mašininio mokymosi algoritmams profiliuojant asmenis, reitinguojant komercinius pasiūlymus, taip padedant priimti algoritmams sprendimus, pakeičiant asmens veikimą.

Darbe analizuojama mašininio mokymosi technologijos samprata ir veikimo jos būdai bei šios technologijos keliamos rizikos. Nagrinėjami pagrindiniai mašininio mokymosi elementai – duomenys ir algoritmai, bei pagrindiniai iššūkiai, su kuriais susiduriama arba, kurie kelia didžiausią riziką asmenų atžvilgiu. Vadovaujantis įvardintomis rizikomis, darbe dėmesys skiriamas trims sritims, kuriose mašininio mokymosi sprendimai turi arba gali turėti teisinį poveikį. Mašininio mokymosi sprendimų poveikis asmens privatumui nagrinėjamas per asmens duomenų apsaugos užtikrinimą ir asmens teisę pačiam nuspręsti dėl asmens duomenų naudojimo. Vartotojų teisių apsaugos problematika nagrinėjama atsižvelgiant į naują reguliavimą Europos Sąjungos mastu, atkreipiant dėmesį į profiliavimo ir reitingavimo keliamas rizikas asmens informuotam pasirinkimui. Atsakomybės už mašininio mokymosi sprendimais atsiradusios žalos klausimai nagrinėjami per doktrinoje keliamus klausimus, dėl būtinųjų civilinės atsakomybės sąlygų identifikavimo ir įrodymo ypatumus. Atlikta analizė identifikuoja svarbius klausimus ir galimus jų sprendimo būdus, kurie turės įtakos besiformuojančiai teismų praktikai mašininio mokymosi sprendimų srityje.

SUMMARY

Most Important Current Legal Challenges of Machine Learning Based Solutions

Gytautė Peseckaitė-Kibickienė

The legal issues analyzed in the master 's thesis are related to machine learning solutions, which are distinguished by the prevailing political discourse and the doctrine of law. The most important current legal challenges are addressed through the elements of personal privacy protection, consumer protection, and liability for damage to machine learning algorithms in profiling individuals, ranking commercial offers, and thus helping algorithms make decisions that alter human behavior.

The paper analyzes the concept of machine learning technology and its modes of operation and the risks posed by this technology. The main elements of machine learning data and algorithms, as well as the main challenges that face or pose the greatest risk to individuals are examined. Based on the identified risks, the work focuses on three areas where machine learning decisions have or may have legal implications. The impact of machine learning decisions on personal privacy is addressed through the protection of personal data and the individual's right to decide on the use of personal data. The issue of consumer protection is being addressed in the context of the new EU-wide regulation, focusing on the risks of profiling and ranking for individuals informed choices. The issues of liability for damage caused by machine learning decisions are addressed through the questions raised in the doctrine regarding the identification and proof of the essential conditions for civil liability. The analysis identifies important issues and possible solutions that will influence the emerging case law in the field of machine learning solutions.