

TARTU ÜLIKOOL
ÕIGUSTEADUSKOND
Avaliku õiguse osakond

Katre Kivi

**EUROOPA LIIDU ISIKUANDMETE ÜLDMÄÄRUSE ISIKUANDMETE
MINIMAALSUSE JA ANONÜMISEERIMISE PÕHIMÕTETE KOHALDAMINE
SUURTELE KEELEMUDELITELE**

Magistritöö

Juhendaja
dr. iur. Mario Rosentau

Tartu
2026

SISUKORD

SISSEJUHATUS

I PEATÜKK

TEHISINTELLEKT JA GDPR-I KOHALDUMINE ISIKUANDMETE TÖÖTLEMISEL

- 1.1. Tehisintellekti ja suurte keelemudelite (LLM-ide) eripära andmetöötles
- 1.2. Isikuandmete määratlus ja GDPR-i kohaldumisa AI-kontekstis
- 1.3. Vastutava ja volitatud töötaja rollid AI arendamisel
- 1.4. Läbipaistvuse ja vastutuse põhimõtted tehisintellekti andmetöötles

II PEATÜKK

ANDMETE MINIMAALSUSE PÕHIMÕTE JA LÄBIPAISTVUSE PROBLEEM SUURTE KEELEMUDELITE ARENDAMISEL

- 2.1. Andmete minimaalsuse põhimõtte sisu ja eesmärk GDPR-i alusel
- 2.2. Andmete minimaalsuse ja eesmärgipärasuse suhe AI-andmetöötles
- 2.3. LLM-ide treeningpraktikad ja andmekogumise läbipaistmatus
- 2.4. Proportsionaalsuse hindamine ja järelevalvepraktika
- 2.5. Tehnilised ja korralduslikud meetmed andmete minimaalsuse ja läbipaistvuse toetamiseks

III PEATÜKK

ANONÜMISEERIMINE JA PSEUDONÜMISEERIMINE TEHISINTELLEKTI RAKENDUSTES

- 3.1. Anonümiseerimise ja pseudonümiseerimise õiguslik eristus GDPR-is
- 3.2. „Pöördumatus“ nõue ja taasidentifitseerimise risk AI kontekstis
- 3.3. Pseudonümiseeritud andmete kasutamise õiguslikud piirid
- 3.4. Anonümiseerimise roll andmete minimaalsuse alternatiivina
- 3.5. Õiguslikud ja praktilised probleemid LLM-ide treeningandmetes

KOKKUVÕTE

SUMMARY

Kasutatud allikad

SISSEJUHATUS

1. Uurimisprobleemi püstitus ja töö eesmärk

Tehisintellekti, eeskätt suurte keelemudelite (*large language models* – LLM), kiire areng on oluliselt muutnud isikuandmete töötlemise ulatust ja viise. Kaasaegsete LLM-ide arendamine ja täiustamine eeldab mahukate ja mitmekesiste andmestike kasutamist, mis võivad sisaldada ka isikuandmeid. Selline andmetöötluspraktika tõstatab küsimuse, kuivõrd on Euroopa Liidu isikuandmete kaitse üldmääruses (GDPR) sätestatud andmekaitsepõhimõtted, eeskätt andmete minimaalsuse põhimõte, sellises kontekstis tegelikult rakendatavad.

GDPR art 5 lg 1 p c kohaselt peavad isikuandmed olema asjakohased, olulised eesmärgi suhtes ja piirduma sellega, mis on vajalik töötlemise eesmärkide saavutamiseks. LLM-ide arendamisel (treenimisel) kasutatakse mahukaid andmekogumeid, ilma et arendajal oleks alati selge ja kontrollitav ülevaade andmete täpsest sisust, päritolust ja isikuandmete ulatusest. See tekitab pinge andmete minimaalsuse normatiivse nõude ja tehisintellekti arenduspraktikate vahel.

Lisaks minimaalsuse põhimõttele LLM-de treenimisel on teiseks keskseks probleemiks anonümiseerimise ja pseudonümiseerimise roll tehisintellekti teostatud andmetöötluses. Nimelt seisneb probleem selles, et nii tehisintellekti arendamisel (nt LLM-ide treenimisel) kui ka AI (*Artificial Intelligence*) abil andmete töötlemisel esineb kõrgendatud risk, et isikuandmeid töödeldakse ebapiisava anonümiseerimise või pseudonümiseerimise tõttu viisil, mis võib kaasa tuua GDPR-i nõuete rikkumise, sealhulgas isiku tuvastatavuse säilimise või taastumise. Kuigi anonümiseerimist käsitletakse praktikas sageli viisina, mis võib vähendada GDPR-ist tulenevate kohustuste ulatust, seab tehisintellekti võimekus andmeid kombineerida ja mustreid tuvastada kõrgemad ootused anonümiseerimise pöördumatusele ning seega ka andmete tegelikule anonüümsusele. Pseudonümiseeritud andmete kasutamise korral säilib vastutaval töötlejal kohustus järgida GDPR-i nõudeid ning tagada jätkuv riskihindamine ja asjakohaste kaitsemeetmete rakendamine.

Uurimisprobleemi teravdab asjaolu, et LLM-ide arendamine ja kasutamine on andmekaitse seisukohalt eristatavad ning erineva läbipaistvuse astmega tegevused. Ühelt poolt on mudelite treenimine ja edasiarendamine (sealhulgas peenhäälestus) sageli piiratud läbipaistvusega, kuna treeningandmete koosseis, allikad ja töötlemisetapid võivad olla raskesti tuvastatavad nii

tehnilise keerukuse kui ka ärisaladuse kaitse tõttu. Teiselt poolt avaldub nn „musta kasti“ probleem eeskätt mudelite kasutamisel, kus sisendite ja väljundite vaheline seos ei ole täielikult selgitatav. Selline olukord raskendab nii järelevalveasutuste kui ka andmesubjektide võimalusi hinnata, kas ja millisel määral järgitakse GDPR-i põhimõtteid.

Käesoleva magistritöö eesmärk on analüüsida, millisel määral ja millistel tingimustel on võimalik ühildada GDPR-ist tulenevad andmete minimaalsuse ning anonümiseerimise ja pseudonümiseerimise nõuded ühelt poolt suurte keelemudelite arendamisega (sealhulgas treenimise ja peenhäälestusega) ning teiselt poolt nende kasutamise ja rakendamisega, samuti tuvastada peamised õiguslikud ja praktilised piirangud, mis sellist ühildamist mõjutavad.

2. Uurimisküsimused ja metoodika

Käesolev magistritöö lähtub eeldusest, et LLM-ide arendamine ja kasutamine seab väljakutseid GDPR-ist tulenevate andmekaitsepõhimõtete, eelkõige andmete minimaalsuse ning anonümiseerimise ja pseudonümiseerimise nõuete kohaldamisele. Sellest tulenevalt otsitakse vastust küsimusele, kuidas on võimalik ühildada andmete minimaalsuse põhimõtte suurte keelemudelite ulatusliku andmevajadusega.

Selle küsimuse täpsustamiseks analüüsitakse, millised tehnilised ja regulatiivsed meetmed võiksid vastata GDPR-i nõudele töödeldavate isikuandmete minimaalsusest suurte keelemudelite arendamisel, ning kuidas kujundavad kehtivad õiguslikud tõlgendused, sealhulgas kohtupraktika ja järelevalveasutuste juhised, minimaalsuse ja proportsionaalsuse hindamist tehisintellekti andmetöötluses.

Lisaks käsitletakse anonümiseerimise ja pseudonümiseerimise rolli tehisintellekti kontekstis, keskendudes küsimusele, milline on nende mõistete õiguslik eristus GDPR-i kohaselt ning millistel tingimustel saab andmeid pidada tegelikult anonüümseks, arvestades tehisintellektiga kaasnevaid taasisentifitseerimisisriske. Samuti hinnatakse, mil määral on kehtivad regulatiivsed tõlgendused piisavad nende riskide adekvaatseks käsitlemiseks.

Käesolevas magistritöös kasutatakse kvalitatiivset ja analüütilist uurimismeetodit, et tagada püstitatud uurimisprobleemi käsitlemine, töö eesmärkide saavutamine ning uurimisküsimustele põhjendatud järeltulemuste tegemine.

Töö metoodiline lähenemine tugineb eeskätt õigusdogmaatilisele analüüsile, mille raames uuritakse GDPR-i asjakohaseid sätteid, eelkõige andmete minimaalsuse, anonümiseerimise ja pseudonümiseerimisega seotud norme, ning nende tõlgendamist Euroopa Kohtu praktikas ja Euroopa andmekaitseasutuste juhistes. Analüüsi eesmärk on selgitada nende põhimõtete normatiivset sisu ning hinnata nende kohaldatavust tehisintellekti, sealhulgas suurte keelemudelite arendamise ja kasutamise kontekstis.

Lisaks rakendatakse töös kriitilist analüüsi, et võrrelda GDPR-ist tulenevaid nõudeid tehisintellekti arendamise tegelike praktikatega ning tuvastada normatiivsete eesmärkide ja tehnoloogiliste reaalsuste vahelisi pingeid. Lõpuks käsitletakse ka järelevalveasutuste seisukohti ja õigusteaduslikke käsitusi, et tuua esile erinevaid tõlgendusviise ning hinnata nende mõju andmekaitsepõhimõtete rakendamisele. Tehnilisi aspekte käsitletakse töös piiratud ulatuses ja üksnes sellisel määral, mis on vajalik õiguslike probleemide mõistmiseks ja analüüsimiseks.

3. Töö struktuur

Magistritöö koosneb sissejuhatusest, kolmest peatükist ning kokkuvõttest. Esimeses peatükis käsitletakse tehisintellekti ja suurte keelemudelite andmetöötluse eripära ning GDPR-i kohaldumist sellises kontekstis, luues töö edasiseks analüüsiks vajaliku õigusliku ja tehnoloogilise raamistiku. Teises peatükis analüüsitakse andmete minimaalsuse põhimõtte sisu ja tähendust suurte keelemudelite arendamisel, pöörates erilist tähelepanu treeningandmete läbipaistvuse probleemile, proportsionaalsuse hindamisele ning võimalikele tehnilistele ja korralduslikele meetmetele minimaalsuse tagamiseks. Kolmandas peatükis keskendutakse anonümiseerimise ja pseudonümiseerimise õiguslikele ja praktilistele piiridele tehisintellekti rakendustes, analüüsides taasidentifitseerimise riske ning anonümiseerimise rolli andmete minimaalsuse alternatiivina. Töö lõpeb kokkuvõttega, milles esitatakse vastused uurimisküsimustele ning tehakse järeldused ja ettepanekud GDPR-i põhimõtete edasiseks rakendamiseks tehisintellekti kontekstis.

TEHISINTELLEKT JA GDPR-I KOHALDUMINE ISIKUANDMETE TÖÖTLEMISEL

1.1. Tehisintellekti ja suurte keelemudelite (LLM-ide) eripära andmetöötluses

Tehisintellekti ehk AI all mõistetakse inimeste loodud masinpõhiseid süsteeme, mis on kavandatud tegutsema teatava autonoomsuse astmega ning saavutama etteantud või kaudselt kujunenud eesmärged füüsilises või digitaalses keskkonnas. Sellised süsteemid võtavad sisendina vastu andmeid, tõlgendavad neid ning teevad järeldusi selle kohta, kuidas genereerida väljundeid – näiteks prognoose, informatsiooni, soovitusi või otsuseid –, mis võivad mõjutada ümbristavat keskkonda.¹ AI-süsteemide eripäraks on nende võime tegutseda osaliselt iseseisvalt ning vajaduse korral kohandada oma toimimist ka pärast kasutuselevõttu, sealhulgas õppida kogemuse põhjal ja muuta oma käitumist vastavalt varasemate tegevuste mõjule.

Rahvusvahelisel tasandil on tehisintellekti määratlemisel oluliseks orientiiriks Organisation for Economic Co-operation and Development (OECD, Majanduskoostöö ja Arengu Organisatsioon) käsitus, mille kohaselt on AI-süsteem masinpõhine süsteem, mis suudab selgesõnaliste või kaudsete eesmärkide alusel järeldada sisendandmetest, kuidas luua väljundeid, mis mõjutavad füüsilisi või virtuaalseid keskkondi. OECD osutab, et AI-süsteemide autonoomia ja kohanemisevõime võivad erineda ning et õppimisvõime pärast kasutuselevõttu ei ole küll alati vajalik, kuid on paljude kaasaegsete AI-rakenduste puhul keskne tunnus. OECD määratlus on kujunenud rahvusvaheliseks standardiks ning on mõjutanud ka Euroopa Liidu tehisintellekti reguleerivat raamistikku.²

Euroopa Liidu (EL) õiguses on tehisintellekti mõistet täpsustatud tehisintellekti määruuses (AI Act), mille kohaselt on AI-süsteem masinpõhine süsteem, mis on loodud töötama erineva autonoomsuse astmega, võib pärast kasutuselevõttu olla kohanduv ning mis suudab sisendite põhjal järeldada, kuidas genereerida väljundeid, mis mõjutavad füüsilisi või virtuaalseid

¹ OECD, *Explanatory Memorandum on the Updated OECD Definition of an AI System*, 2023; uuendatud 2024, kättesaadav: https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html.

² OECD, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449), algselt vastu võetud 22.05.2019, muudetud 08.11.2023 ja 03.05.2024, kättesaadav: <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

keskkondi.³ EL-i määratluses on olulisel kohal just järelduslik iseloom (*inference*), mille kaudu AI-süsteemid eristuvad tavapärasest tarkvarast, ning elutsükli põhine vaade, mis hõlmab nii süsteemi arendamist, kasutuselevõttu kui ka hilisemat kohanemist ja kasutamist. EL-i määratluses on olulisel kohal järelduslikkus, mis eristab AI-süsteeme tavapärasest tarkvarast, ning elutsükli põhine lähenemine, mis hõlmab süsteemi arendamist, kasutuselevõttu ning selle hilisemat toimimist ja võimalikku kohandumist.

Teadusdistsipliinina hõlmab tehisintellekt mitmekesist meetodite ja käsitusviiside kogumit, nagu masinõpe (sealhulgas süvaõpe ja tugevdusõpe), mille puhul süsteemid õpivad mustreid treeningandmetest, masinarutlus, mis hõlmab planeerimist, otsingut, optimeerimist ning teadmiste esitust ja järeldamist, ning robotika, kus AI-süsteemid on lõimitud füüsiliste seadmetega.⁴ Need lähenemised võivad toimida eraldiseisvalt või kombineeritult ning moodustavad aluse kaasaegsetele AI-rakendustele, sealhulgas suurtele keelemudelitele.

LLM-ide kasutuselevõtt on muutnud eeskätt tekstipõhiste vestlusmodelite arendamist ja kasutamist ning tõstatanud sellega seoses uusi tehnilisi ja õiguslikke küsimusi. LLM-id põhinevad süvaõppe meetoditel, eelkõige nn transformer-arhitektuuril, mis võimaldab töödelda ja genereerida loomulikku keelt ulatuslike treeningandmetike põhjal. Sellised mudelid ei tugine etteantud reeglitele, vaid õpivad keelelisi mustreid suurtest tekstikorpustest, võimaldades täita mitmesuguseid ülesandeid alates tekstiloomest kuni keerukate analüütiliste funktsioonideni.⁵

LLM-ide andmetöötluse eripära seisneb eeskätt nende mastaabis ja andmete heterogeensuses. Mudelite treenimiseks kasutatakse tavaliselt väga suuri ja mitmekesiseid andmekogumeid, mis võivad sisaldada nii struktureeritud kui ka struktureerimata andmeid, sealhulgas veebitekste, raamatuid, artikleid ja kasutajate loodud sisu. Selline lähenemine tähendab, et treeningandmete

³ Euroopa Parlamendi ja nõukogu määrus (EL) 2024/1689, 13. juuni 2024, millega kehtestatakse tehisintellekti käsitlevad ühtlustatud normid (tehisintellekti määrus), ELT L 2024/1689, 12.7.2024 Lõige 12 sätestab, et „Tehisintellekti süsteemi mõiste tuleks käesolevas määruses selgelt määratleda ning see peaks olema tihedalt kooskõlas tehisintellekti valdkonnas tegutsevate rahvusvaheliste organisatsioonide tööga, et tagada õiguskindlus, hõlbustada rahvusvahelist ühtlustamist ja laialdast aktspteerimist, võimaldades samal ajal paindlikkust, et kohaneda selle valdkonna kiire tehnoloogilise arenguga“

⁴ OECD, *Explanatory Memorandum on the Updated OECD Definition of an AI System*, OECD Publishing, 2023/2024, eriti jaotised AI-süsteemide autonoomia, järelduslikkuse ja kohanemise võime kohta, veebist kättesaadav: https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html

⁵ Bender, E. M. (2024). *The Social Impact of Large Language Models*. *Communications of the ACM*, 67(1). <https://doi.org/10.1145/3639372>

täpne koosseis ei ole sageli täielikult teada ega kontrollitav, mis omakorda raskendab hinnangut selle kohta, millises ulatuses sisaldavad need isikuandmeid või muid tundlikke andmeid.⁶

Oluline eripära on ka LLM-ide õppimisviis ja teadmiste representatsioon. Mudelid ei salvesta andmeid traditsioonilises andmebaasilises tähenduses, vaid kodeerivad mustreid ja seoseid oma parameetrites. Sellest hoolimata on teaduskirjanduses järeltatu, et mudelid võivad teatud juhtudel meelde jätta treeningandmete fragmente ning neid hiljem väljundina taastoota. Andmelekkeks (data leakage) kirjeldatav olukord võib viia selleni, et mudel väljastab isikuandmeid ka ilma otsese juurdepääsuta algandmetele.⁷

Lisaks iseloomustab LLM-e mitte-deterministlikkus, mis tähendab, et sama sisendi korral ei anna mudel identset vastust. See tuleneb nii mudelite arhitektuurist kui ka kasutatavatest juhuslikkuse komponentidest (nt sampling-meetodid). Sampling-meetodite puhul valib mudel järgmise sõna mitme võimaliku variandi hulgast juhuslikkust arvestades, mistõttu võivad vastused varieeruda. Selline omadus raskendab andmetöötamise prognoositavust ja kontrollitavust, mis on aga klassikalises andmekaitseõiguses keskse tähtsusega.

Teine oluline aspekt on LLM-ide läbipaistmatus, mida käsitletakse sageli mudeli piiratud selgitatavuse probleemina. Kuigi mudelid võivad anda sisuliselt korrektseid vastuseid, ei ole nende toimeloogika sageli piisavalt selgitatav. Nagu on märkinud Euroopa Andmekaitsekoostöögrupp, jäävad LLM-ide sisemised mehhanismid suuresti raskesti selgitatavaks ning nende mõistmine eeldab täiendavate interpretatsioonimeetodite kasutamist.⁸ See tekitab probleeme mitte ainult tehnilises mõttes, vaid ka isikuandmete kaitse perspektiivist, kus läbipaistvus ja põhjendatavus on olulised põhimõtted.

LLM-ide andmetöötamise eripära avaldub kogu selle elutsükli ulatuses. Andmetöötlus ei piirdu üksnes mudeli treenimisega, vaid hõlmab ka:

⁶ Feretzakis, G. *et al.*, *GDPR and Large Language Models: Technical and Legal Obstacles*, *Future Internet* 17(4), 151 (2025), <https://doi.org/10.3390/fi17040151>

⁷ Ran Zhang, Hong-Wei Li, Xin-Yuan Qian, Wen-Bo Jiang, Han-Xiao Chen. On large language models safety, security, and privacy: A survey[J]. *Journal of Electronic Science and Technology*, 2025, 23(1): 100301. <https://www.researching.cn/articles/OJba348e2553344135>

⁸ European Data Protection Board, *AI Privacy Risks & Mitigations – Large Language Models (LLMs)*, Support Pool of Experts Programme (2025), kättesaadav: <https://www.edpb.europa.eu/system/files/2025-04/ai-privacy-risks-and-mitigations-in-llms.pdf>

1. andmete kogumist ja eelprotsessimist,
2. mudeli treenimist ja valideerimist,
3. mudeli juurutamist ja kasutamist,
4. pidevat täiustamist ja uuendamist.

Igas etapis võivad tekkida erinevad andmekaitseriskid, sealhulgas volitamata andmete kasutamine, andmete lekke oht ning raskused andmete kustutamisel või parandamisel pärast mudeli treenimist.⁹

Siinkohal saab järeldada, et LLM-ide andmetöötlus erineb oluliselt traditsioonilisest andmetöötlusest. Seda iseloomustavad suur andmemaht, piiratud kontroll treeningandmete üle, mudelite võime tuletada ja taastoota infot ning läbipaistvuse puudumine. Need omadused loovad olulise lähtekoha järgmistes alapeatükkides käsitletavatele küsimustele, eelkõige seoses isikuandmete määratluse, GDPR-i kohaldumisala ning vastutuse ja läbipaistvuse põhimõtete rakendamisega tehisintellekti kontekstis.

1.2. Isikuandmete mõiste ja GDPR-i kohaldumisala AI-kontekstis

GDPR-i kohaldatavus tehisintellekti, eelkõige LLM-ide kontekstis sõltub esmajoones sellest, kas ja millises ulatuses käsitletavat andmed kvalifitseeruvad isikuandmeteks. GDPR artikli 4 lõike 1 kohaselt on isikuandmed igasugune teave tuvastatud või tuvastatava füüsilise isiku kohta. Kuigi definitsioon on tehnoloogianeutraalne ja laiaulatuslik, tekitab selle rakendamine AI-kontekstis mitmeid tõlgenduslikke probleeme, mis tulenevad eeskätt andmete töötlemise eripärast ja mudelite toimimisloogikast.

Traditsioonilises andmetöötluses on isikuandmete tuvastamine suhteliselt selgepiiriline, kuna andmed paiknevad struktureeritud kujul ja on otseselt seotud konkreetsete isikutega. LLM-ide puhul on olukord märksa keerukam. Treeningandmestikud sisaldavad sageli massiliselt veebist kogutud tekste, mille hulka võivad kuuluda nii avalikult kättesaadavad kui ka potentsiaalselt isikuandmeid sisaldavad andmed. Seejuures ei ole arendajatel sageli täielikku ülevaadet andmestike täpsest sisust ega päritolust, mis muudab isikuandmete olemasolu hindamise

⁹ *Ibid.*

ebakindlaks.¹⁰ Selline määramatus seab kahtluse alla klassikalise eelduse, et andmetöötleva suudab alati määratleda, milliseid andmeid ta töötleb.

Lisaks kerkib küsimus sellest, kas ja millal võib LLM-i mudel ise sisaldada isikuandmeid. Kuigi mudelid ei salvesta andmeid otseselt, on uuringud näidanud, et need võivad teatud juhtudel taastoota treeningandmete fragmente, sealhulgas isikuandmeid.¹¹ Eeltoodu viitab sellele, et isikuandmed võivad olla kaudselt kodeeritud mudeli parameetrites. Selline nähtus laiendab traditsioonilist arusaama andmetöötlemisest, kuna isikuandmete olemasolu ei pruugi olla tuvastatav ilma spetsiifiliste tehniliste meetoditeta. Sellest tulenevalt võib väita, et ka mudeli enda käsitlemine pelgalt anonüümse tootena ei pruugi alati olla õigustatud.

Oluline küsimus tõusetub ka LLM-ide väljundite hindamisel. Kui mudel genereerib teksti, mis sisaldab tuvastatava isiku kohta käivat teavet, tuleb hinnata, kas tegemist on isikuandmete töötlemisega GDPR-i tähenduses. Euroopa Kohtu praktika kohaselt tuleb isikuandmete mõistet tõlgendada laialt ning arvesse tuleb võtta kõiki vahendeid, mida on mõistlikult võimalik kasutada isiku tuvastamiseks (vt nt Breyer vs Saksamaa, C-582/14).¹² Sellest vaatenurgast võib ka LLM-i genereeritud sisu kvalifitseeruda isikuandmeteks, kui see võimaldab isikut otseselt või kaudselt tuvastada. Samas tekitab see küsimuse vastutuse piiridest olukorras, kus väljund ei ole otseselt eelnevalt salvestatud andmete koopia, vaid statistiline tuletis.

GDPR-i kohaldumisala määratlemisel on küsimus, kas LLM-ide arendamine ja kasutamine kujutab endast isikuandmete töötlemist määruse artikli 4 lõike 2 tähenduses. Arvestades, et töötlemine hõlmab mh andmete kogumist, säilitamist, kasutamist ja avalikustamist, on selge, et vähemalt treeningfaasis toimub andmete töötlemine. Keerulisemaks muutub olukord mudeli kasutusfaasis, kus sisend- ja väljundandmete roll ei ole alati üheselt määratletav. Mõned autorid on väitnud, et ka pelgalt mudeli kasutamine võib kujutada endast andmetöötlust, kui see hõlmab isikuandmete sisestamist või genereerimist.¹³

¹⁰ Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., *et al.* (2021). *On the opportunities and risks of foundation models*. arXiv preprint arXiv:2108.07258. Kättesaadav: <https://arxiv.org/abs/2108.07258>

¹¹ Carlini, N., Tramer, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., *et al.* (2021). *Extracting training data from large language models*. Proceedings of the 30th USENIX Security Symposium, 2633–2650. Kättesaadav: <https://arxiv.org/abs/2012.07805>

¹² Court of Justice of the European Union. (2016). *Patrick Breyer v Bundesrepublik Deutschland* (C-582/14). *ECLI: ECLI:EU:C:2016:779* Kättesaadav: <https://curia.europa.eu/juris/liste.jsf?num=C-582/14>

¹³ Veale, M., & Borgesius, F. Z. (2021). *Demystifying the draft EU Artificial Intelligence Act*. *Computer Law Review International*, 22(4), 97–112. <https://doi.org/10.9785/cr-2021-220402>

Siinkohal ilmneb regulatiivne vastuolu: ühelt poolt on GDPR-i eesmärk tagada isikute kõrgetasemeline kaitse, teisalt tuleb põhimõtete kohaldamisel arvestada tehisintellekti süsteemide tehniliste eripärade ja kasutusjuhtude mitmekesisusega. See eeldab kontekstipõhist hinnangut, eeskätt siis, kui isikuandmete töötlemine on süsteemi toimimiseks vältimatu või kui süsteemi kasutamisel võib olla oluline mõju andmesubjektile. Selline lähenemine ei tähenda andmekaitse nõuete lõdvendamist, vaid proportsionaalsuse ja riskihindamise sisulist rakendamist.

Praktikas tähendab seda, et suurema mõjuga kasutusjuhtudel, kus süsteemi väljund mõjutab andmesubjekti õigusi või olulisi huve, peavad õiguspärasuse, läbipaistvuse ja riskimaandamise nõuded olema rangemalt täidetud. Väiksema mõjuga kasutusjuhtudel ei kao GDPR-i põhimõtted samuti, kuid nende sisuline rakendamine võib erineda, arvestades töötlemise eesmärki, andmete liike ja võimalikke tagajärgi. Sellest tulenevalt ei ole andmekaitse ja innovatsioon käsitatavad vastanditena, vaid küsimus taandub õiguspärase ja proportsionaalse töötlemise tingimuste määratlemisele.

Kahtlemata saab järeldada, et GDPR-i kohaldumine AI-kontekstis ei ole mitte niivõrd küsimus selle formaalsest ulatusest, vaid pigem selle sisulisest tõlgendamisest uutes tehnoloogilistes tingimustes. Isikuandmete mõiste jääb küll laiaulatuslikuks ja paindlikuks, kuid selle rakendamine LLM-ide puhul nõuab täiendavat analüüsi, mis arvestab mudelite tehnilist eripära, andmete kaudset esitust ning väljundite ettearvatust. See loob aluse järgmistes alapeatükkides käsitletavale vastutuse ja läbipaistvuse probleemistikule.

1.3. Vastutava ja volitatud töötleja rollid AI arendamisel

Isikuandmete töötlemise õiguspärasus tehisintellekti kontekstis sõltub olulisel määral sellest, kuidas määratletakse vastutava ja volitatud töötleja rollid. GDPR artikli 4 lõigete 7 ja 8 kohaselt on vastutav töötleja see, kes määrab kindlaks isikuandmete töötlemise eesmärgid ja vahendid, samas kui volitatud töötleja tegutseb vastutava töötleja nimel ja juhiste alusel. Näiteks kui ettevõtte kasutab vestlusrobotit kliendipäringutele vastamiseks, siis ettevõtte ise on vastutav töötleja, kuna otsustab, miks ja kuidas klientide andmeid kasutatakse. Vestlusroboti

teenusepakkuja on volitatud töötleja, sest ta töötleb andmeid ettevõtte juhiste järgi. Kuigi selline eristus on traditsioonilises andmetöötluses üldjuhul selge, muutub see tehisintellekti, eriti suurte keelemudelite puhul märkimisväärselt keerukamaks.

LLM-ide arendamine ja kasutamine toimub tavaliselt mitmetasandilises ökosüsteemis, kus osalevad erinevad subjektid, sealhulgas mudeli arendajad, platvormipakkujad ning lõppkasutajad. Näiteks võib mudeli arendaja koguda ja töödelda treeningandmeid, samas kui teenusepakkuja integreerib mudeli oma süsteemi ning lõppkasutaja kasutab seda konkreetsete ülesannete täitmiseks. Sellises ahelas ei ole alati üheselt selge, kes määrab töötlemise eesmärgid ja vahendid, mis on aga vastutava töötleja määratlemise kriteerium.¹⁴

AI-süsteemide puhul võib esineda olukordi, kus mitu osapoolt kvalifitseeruvad samaaegselt vastutavateks töötlejateks ehk nn ühine vastutus (*joint controllership*).¹⁵ Näiteks juhul, kui nii mudeli arendaja kui ka selle rakendaja mõjutavad oluliselt seda, kuidas ja milleks andmeid kasutatakse, võib olla põhjendatud käsitleda neid ühiste vastutavate töötlejatena. Samas ei ole selline lähenemine alati praktiline, kuna vastutuse jagamine võib muutuda ebaselgeks ning raskendada andmesubjektide õiguste teostamist.

Täiendav keerukus tuleneb sellest, et LLM-id on sageli välja töötatud üldotstarbeliste süsteemidena, mille konkreetne kasutusviis ei ole arendamise hetkel täielikult ette nähtav. See tähendab, et arendaja ei pruugi määrata lõplikke andmetöötluse eesmärgi, vaid need kujunevad alles kasutusfaasis. Sellises olukorras tekib küsimus, kas arendajat saab käsitada vastutava töötlejana või pigem tehnoloogilise vahendi pakkujana. Mõned autorid on leidnud, et arendaja roll võib sarnaneda tootja rolliga, kes ei ole otseselt vastutav iga konkreetse kasutusjuhtumi eest.¹⁶ Teisalt ei saa täielikult välistada arendaja vastutust, eriti juhul, kui ta määrab olulised tehnilised parameetrid, mis mõjutavad andmetöötluse olemust.

Volitatud töötleja rolli määratlemine on AI kontekstis samuti problemaatiline. Klassikalises tähenduses eeldab volitatud töötleja, et ta tegutseb selgelt määratletud juhiste alusel. LLM-ide

¹⁴ Veale, M., & Borgesius, F. Z., *Demystifying the draft EU Artificial Intelligence Act*, *Computer Law Review International* 22(4), 97–112 (2021), <https://doi.org/10.9785/crl-2021-220402>.

¹⁵ Wachter, S., Mittelstadt, B., & Floridi, L., *Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation*, *International Data Privacy Law* 7(2), 76–99 (2017), <https://doi.org/10.1093/idpl/ix005>

¹⁶ Kuner, C., Bygrave, L. A., & Docksey, C. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary*, Oxford University Press, 2020.

puhul ei pruugi aga selline hierarhiline suhe alati eksisteerida. Näiteks pilvepõhise AI-teenuse pakkuja võib küll töödelda andmeid kliendi nimel, kuid samal ajal mõjutada oluliselt töötlemise tehnilisi vahendeid ja isegi eesmäärke. Sellisel juhul võib tekkida olukord, kus formaalselt volitatud töötleja rollis olev subjekt omab sisuliselt suuremat kontrolli, kui GDPR-i klassikaline mudel eeldab. Kirjeldatud rolli ebamäärasustest tulenevaid riske aitab küll maandada põhjalik ja selge lepinguline regulatsioon, milles on täpselt määratletud poolte vastutus, rollid ja andmetöötluse tingimused, kuid pelgalt leping ei pruugi alati kõrvaldada olukordi, kus tegelik kontroll ja otsustusõigus kaldub GDPR mõttes volitatud töötleja kasuks.

Saan nõustuda, et GDPR-i traditsiooniline rollijaotus ei pruugi olla piisav tehisintellekti ökosüsteemi kirjeldamiseks. Eelkõige on probleemiks see, et regulatsioon eeldab suhteliselt lineaarset ja hierarhilist andmetöötluse mudelit, samas kui AI-süsteemid toimivad sageli hajusates ja dünaamilistes võrgustikes. Sellest tulenevalt võiks kaaluda paindlikumat lähenemist, mis arvestab erinevate osapoolte tegelikku mõju andmetöötlusele, mitte üksnes formaalseid rolle. Eluliselt tähendab see näiteks olukordi, kus AI-süsteem koosneb mitmest osapooltest – üks arendab mudeli, teine pakub pilveteenust ja kolmas kasutab seda oma klientide jaoks, mistõttu on raske öelda, kes tegelikult andmete töötlemist kontrollib. Samuti võivad süsteemid aja jooksul kasutajate sisenditest õppida nii, et rollid muutuvad ja hägustuvad. Avatud lähtekoodiga mudelite puhul ei ole üldse üht selget vastutajat ning andmed liiguvad sageli läbi mitme tehnilise kihi, kus iga osapool mõjutab tulemust. Sellised olukorrad näitavad selgelt, et tegelik kontroll ei pruugi vastata formaalsetele rollidele, mida regulatsioon eeldab.

Andmesubjekti jaoks ei ole esmatähtis, milline osapool on formaalselt vastutav või volitatud töötleja, vaid see, kas tema õigused on tegelikult kaitstud ja teostatavad. Kui vastutuse jaotus on ebaselge, võib see viia olukorrani, kus ükski osapool ei võta täielikku vastutust, mis omakorda nõrgendab GDPR-i eesmärki tagada isikute tõhus kaitse. See omakorda loob vajaduse kas olemasolevate mõistete paindlikumaks tõlgendamiseks või uute regulatiivsete lähenemiste väljatöötamiseks, mis suudaksid paremini arvestada tehisintellekti tehnoloogilise reaalsusega.

1.4. Läbipaistvuse ja vastutuse põhimõtted tehisintellekti andmetöötluses

Läbipaistvuse ja vastutuse põhimõtted on GDPR-i olulise tähtsusega alustalad, mis omavad tehisintellekti kontekstis erilise kaalu. GDPR artikli 5 lõike 1 punkt a sätestab läbipaistvuse

põhimõtte, mille kohaselt peab isikuandmete töötlemine olema andmesubjekti jaoks arusaadav ja jälgitav, samas kui artikli 5 lõige 2 kehtestab vastutuse põhimõtte (*accountability*), mille kohaselt peab vastutav töötleja olema suuteline tõendama, et andmetöötlus vastab kehtivatele nõuetele. Kuigi need põhimõtted on tehnoloogianeutraalsed, seab nende rakendamine LLM-ide puhul mitmeid sisulisi ja praktilisi väljakutseid.

Läbipaistvuse nõue eeldab, et andmesubjektile on võimalik selgitada, kuidas ja millisel eesmärgil tema andmeid töödeldakse. Tehisintellekti süsteemide puhul, eriti süvaõppel põhinevate mudelite korral, on see aga sageli raskendatud nende keerukuse ja nn musta kasti (*black box*) olemuse tõttu. LLM-id töötlevad andmeid viisil, mis ei ole otseselt inimesele vaadeldav, mistõttu sageli ei ole võimalik anda selget ja arusaadavat selgitust konkreetse väljundi kujunemise kohta.¹⁷ See tekitab kahtluse, kas traditsiooniline arusaam läbipaistvusest – st selge, arusaadav ja konkreetsele isikule suunatud teave – on tehisintellekti kontekstis üldse täiel määral realiseeritav.

Läbipaistvust ei saa käsitada üksnes tehnilise seletatavuse (*explainability*) küsimusena, vaid see hõlmab ka laiemat institutsionaalset ja protseduurilist mõõdet.¹⁸ See tähendab, et isegi juhul, kui üksiku otsuse loogikat ei ole võimalik täielikult avada, peaksid süsteemi toimimise üldised põhimõtted, andmeallikad ja riskid olema piisavalt selgitatud. LLM-ide puhul on see aga problemaatiline, kuna treeningandmete päritolu ja koostis ei ole sageli täielikult dokumenteeritud või avalikustatud, mis omakorda piirab sisulise läbipaistvuse saavutamist.

Vastutava töötleja kohustus tagada ja tõendada isikuandmete töötlemise vastavust GDPR-i nõuetele süvendab neid probleeme veelgi, kuna see eeldab mitte ainult nõuete järgimist, vaid ka nende järgimise tõendamist. Tehisintellekti süsteemide puhul võib aga olla keeruline näidata, et kõik GDPR-i põhimõtted – nagu eesmärgipärasus, minimaalsus või täpsus – on tegelikult täidetud. Näiteks võib olla praktiliselt võimatu tõendada, et mudeli treenimisel kasutatud andmed ei sisaldanud ebaõigeid või ebaseaduslikult kogutud isikuandmeid. Selline olukord

¹⁷ Jenna Burrell 2016. *How the machine 'thinks': Understanding opacity in machine learning algorithms*. Big Data & Society, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>

¹⁸ Mike Ananny, Kate Crawford 2018. *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*. New Media & Society, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>

ilmestab, et vastutuse põhimõtte rakendamine AI-kontekstis ei ole üksnes juriidiline, vaid ka tehniline probleem.¹⁹

Selgub, et kui traditsioonilises andmetöötluses on rõhk olnud eelkõige tagantjärele kontrollil ja dokumenteerimisel, siis tehisintellekti puhul tuleb suuremat tähelepanu pöörata ennetavale riskijuhtimisele. See tähendab, et läbipaistvuse ja vastutuse tagamine ei saa piirduda üksnes hilisemate selgitustega, vaid peab olema arvestatud juba süsteemi kavandamisel ja arendamisel. Seetõttu muutub oluliseks hinnata võimalikke riske varajases etapis ning kujundada süsteem selliselt, et andmekaitse põhimõtted oleksid sellesse algusest peale sisse ehitatud. Selline lähenemine on kooskõlas ka „privacy by design“ põhimõttega, mille kohaselt tuleb andmekaitse integreerida tehnoloogiasse juba selle loomise faasis. „Privacy by design“ ehk lähteandmekaitse põhimõte tuleneb DGPR artiklist 25.

Samas ei saa mööda vaadata ka sellest, et täielik läbipaistvus võib olla teatud juhtudel vastuolus teiste õigustatud huvidega, näiteks ärisaladuste kaitse või süsteemi turvalisusega. LLM-ide arendajad võivad põhjendatult soovida kaitsta oma mudeleid ja andmestikke konkurentide eest, mis piirab teabe avalikustamise ulatust.²⁰ Seega tuleb läbipaistvuse põhimõtet rakendada tasakaalustatult, arvestades nii andmesubjektide õigusi kui ka innovatsiooni ja majandushuve.

¹⁹ Nicholas Diakopoulos 2016. *Accountability in algorithmic decision making*. Communications of the ACM, 59(2), 56–62. <https://doi.org/10.1145/2844110>

²⁰ Lilian Edwards, Michael Veale 2017. *Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for*. Duke Law & Technology Review, 16, 18–84.

II PEATÜKK

ANDMETE MINIMAALSUSE PÕHIMÕTE JA LÄBIPAISTVUSE PROBLEEM SUURTE KEELEMUDELITE ARENDAMISEL

2.1. Andmete minimaalsuse põhimõtte sisu ja eesmärk GDPR-i alusel

Andmete minimaalsuse põhimõte on üks GDPR-i keskseid andmekaitsepõhimõtteid, mille eesmärk on piirata isikuandmete töötlemist ulatuses, mis on vajalik konkreetse ja õiguspärase eesmärgi saavutamiseks. Vastavalt GDPR artikli 5 lõike 1 punktile c peavad töödeldavad isikuandmed olema asjakohased, olulised ja piiratud sellega, mis on vajalik nende töötlemise eesmärgi seisukohalt. See põhimõte ei kujuta endast üksnes kvantitatiivset piirangut andmete kogumisele, vaid sisaldab ka kvalitatiivset mõõdet, nõudes andmete sisulist seost töötlemise eesmärgiga. Andmete minimaalsuse põhimõte ei piirdu üksnes kogutavate andmete mahu vähendamisega, vaid eeldab andmete sisulist seost töötlemise eesmärgiga, nõudes hinnangut nende asjakohasusele ja vajalikkusele konkreetse andmetöötluse kontekstis.²¹

Traditsiooniliselt on andmete minimaalsuse põhimõtet käsitatud kui mehhanismi, mis vähendab andmesubjektidele tekkivaid riske, sealhulgas andmete väärkasutuse, lekkimise ja eraellu ebaproportsionaalse sekkumise ohtu. Väiksema hulga andmete kogumine tähendab üldjuhul ka väiksemat potentsiaalset kahju, mistõttu minimaalsuse põhimõte toimib ennetava kaitsemehhanismina.²² See lähenemine eeldab aga, et andmetöötleja suudab ette kindlaks määrata, millised andmed on konkreetse eesmärgi saavutamiseks vajalikud ning millised mitte.

Oluline seos ilmneb eesmärgipärasuse põhimõttega (*purpose limitation*), mille kohaselt tuleb isikuandmeid koguda selgelt määratletud ja õiguspärastel eesmärkidel. Andmete minimaalsus on sisuliselt eesmärgipärasuse mõõdetavaks muutmine. Ilma selge eesmärgita ei ole võimalik hinnata, milline andmehulk on vajalik. Need põhimõtted toimivad vastastikku toetavatena – mida täpsemalt on eesmärk määratletud, seda rangemalt saab rakendada minimaalsuse nõuet.²³

²¹ Kaimar Karu, „Tehisintellekti keerukad küsimused“, *Juridica* 2021/1, lk 45, kättesaadav: <https://ojs.utlib.ee/index.php/juridica/article/view/18277>

²² Gellert, R., *Data protection: A risk regulation? Between the risk management of everything and the precautionary alternative*, *International Data Privacy Law* 5(1), 3–19 (2015), <https://doi.org/10.1093/idpl/ipu035>

²³ Kuner, C., Bygrave, L. A., & Docksey, C. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary*, Oxford University Press, 2020, kättesaadav: <https://academic.oup.com/oxford-law-pro/book/41324>

Sellest tuleneb ka oluline järeldus: minimaalsuse põhimõte ei ole autonoomne, vaid sõltub tugevalt sellest, kui konkreetset on andmetöötamise eesmärgid sõnastatud.

Kuigi minimaalsuse põhimõte on normatiivselt selge, muutub selle rakendamine praktikas sageli keerukaks, eriti olukordades, kus andmetöötamise eesmärgid on laiad või dünaamilised. Digitaalsete teenuste ja andmemahukate süsteemide kontekstis on minimaalsuse põhimõte järk-järgult nihkunud ideaaltasemelt praktiliste kompromisside suunas, kus andmete kogumine toimub osaliselt ka tulevaste, veel määratlemata kasutusviiside tarbeks.²⁴ Selline andmete kogumise praktika seab kahtluse alla minimaalsuse põhimõtte algse loogika, kuna võimaldab andmete kogumist ka ilma selgelt määratletud eesmärgita.

Arusaadavalt seisneb andmete minimaalsuse põhimõtte peamine probleem selles, et mida väärtuslikumad ja mitmekülgsemalt kasutatavad andmed on, seda keerulisem on nende kogumist piirata. Eriti ilmneb see teadus- ja arendusandmetike ja uurimistööde puhul, kus andmete tegelik väärtus ei ole kogumise hetkel sageli teada, vaid selgub alles hilisema analüüsi käigus. Sellises olukorras ei ole minimaalsuse põhimõtte enam üheselt mõõdetav, vaid pigem eelnev hinnang, mida tuleb anda piiratud teabe tingimustes.

Lisaks tõstatab minimaalsuse põhimõtte küsimuse selle rakendamise praktilistest kriteeriumidest. GDPR ei anna täpseid juhiseid selle kohta, kuidas hinnata, milline andmehulk on vajalik, jättes selle suuresti andmetöötaja hinnangu ja järelevalveasutuste tõlgenduse hooleks. See loob olukorra, kus minimaalsuse põhimõtte rakendamine võib olla ebaühtlane ning sõltuda konkreetsetest kontekstidest ja praktikatest.²⁵ Selline paindlikkus on küll vajalik tehnoloogilise neutraalsuse tagamiseks, kuid samas võib see vähendada õiguskindlust.

Andmete minimaalsuse põhimõtet ei saa käsitada pelgalt tehnilise reeglina, mis piirab kogutavate andmete hulka, vaid seda tuleb mõista kui normatiivset instrumenti, mille eesmärk on juhtida andmetöötamise loogikat tervikuna. Tehisintellekti ja andmemahukate süsteemide kontekstis ilmneb aga, et see loogika ei ühti enam tehnoloogilise arengu tegelikkusega.

²⁴ Mayer-Schönberger, V., & Cukier, K., *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Houghton Mifflin Harcourt, 2013, kättesaadav: https://openlibrary.org/works/OL19354258W/Big_Data_A_Revolution_That_Will_Transform_How_We_Live_Work_and_Think.

²⁵ Gellert, R., *Data protection: A risk regulation? Between the risk management of everything and the precautionary alternative*, *International Data Privacy Law* 5(1), 3–19 (2015), <https://doi.org/10.1093/idpl/ipu035>.

Andmete kogumine ei lähtu enam üksnes konkreetsest ja kitsalt määratletud eesmärgist, vaid sageli ka potentsiaalsest tulevase kasutusest, mis muudab minimaalsuse hindamise sisuliselt ebakindlaks.²⁶ Oluline järeldus on, et minimaalsuse põhimõtte rakendamine ei saa AI-kontekstis tugineda üksnes andmete kvantitatiivsele vähendamisele. Selle asemel tuleks rõhk suunata sellele, kuidas andmeid kasutatakse, milliseid riske need kaasa toovad ning milliseid tagajärgi võib nende töötlemine andmesubjektidele põhjustada.²⁷ See viitab vajadusele nihutada regulatiivset fookust andmete minimaalsuse põhimõttelt (*data minimisation*) riskide vähendamisele (*risk minimisation*), kus küsimus ei ole mitte kogutavate andmete hulk, vaid nende töötlemise mõju. Lisaks võiks kaaluda minimaalsuse põhimõtte täpsustamist tehisintellekti kontekstis selliselt, et see hõlmaks selgemaid nõudeid andmekogumite kvaliteedile ja päritolu dokumenteerimisele. Kui andmete vähendamine täieliku miinimumini ei ole realistlik, muutub oluliseks, et andmete kasutamine oleks jälgitav ja põhjendatav. Selline lähenemine võimaldaks säilitada minimaalsuse põhimõtte regulatiivse eesmärgi, kohandades selle rakendamist tehnoloogilise reaalsusega.

Sarnaselt eeltooduga leian, et andmete minimaalsuse põhimõtte ei ole tehisintellekti kontekstis oma tähendust kaotanud, kuid selle traditsiooniline tõlgendus vajab sisulist ümbermõtestamist. Tulevikus sõltub selle põhimõtte tõhusus eelkõige sellest, kas suudetakse leida tasakaal tehnoloogilise innovatsiooni ja andmesubjektide õiguste kaitse vahel viisil, mis ei taandu formaalsele vastavusele, vaid tagab ka tegeliku ja sisulise kaitse.

2.2. Andmete minimaalsuse ja eesmärgipärasuse suhe AI-andmetöötluses

Andmete minimaalsuse ja eesmärgipärasuse põhimõtted moodustavad GDPR-i süsteemis lahutamatu terviku, mille kaudu piiratakse isikuandmete töötlemise ulatust ja eesmärke. Õiguslikus plaanis eeldab see, et andmetöötleja määratleb enne andmete kogumist selgelt töötlemise eesmärgi ning kogub üksnes neid andmeid, mis on selle eesmärgi saavutamiseks vajalikud. Kuid LLM-ide kontekstis see eeldus enam suuresti ei kehti.

Euroopa Parlamendi uuringud on osutanud, et tehisintellekti ja suurandmete kasutamine tekitab otsese pinget traditsiooniliste andmekaitsepõhimõtetega, eelkõige eesmärgipärasuse ja

²⁶ Pärn-Lee, Evelin (2019) *Isikuandmete töötlemise eesmärgipärasus ja selle piirid andmepõhises ühiskonnas*. Juridica, 2019/3, lk 175–184

²⁷ Künnapas, K. (2020). *Riskipõhine lähenemine isikuandmete kaitses ja selle piirid*. Juridica, 2020/6, lk 401–409

minimaalsusega, kuna AI toimimine eeldab suurte andmekogumite kasutamist ning sageli ka nende töötlemist eesmärkidel, mis ei ole kogumise hetkel täielikult määratletud.²⁸ Sellest tulenevalt ei ole tegemist üksnes praktilise raskusega, vaid struktuurse vastuoluga kahe erineva regulatiivse loogika vahel. Illustreerivaks eluliseks näiteks on see, et Euroopa Parlamendi IT-osakond on ametlikel tööseadmetel välja lülitanud tehisintellekti funktsioonid, viidates turva- ja privaatsusriskidele.²⁹

Kirjeldatu on kui nihe eesmärgipõhiselt andmekogumiselt andmepõhisele eesmärgistamisele. Kaasaegsetes andmepõhistes süsteemides kujunevad töötlemise eesmärgid sageli alles andmete analüüsi käigus, mistõttu traditsiooniline eesmärgipärasuse põhimõte ei suuda enam täita oma piiravat funktsiooni.³⁰ See tähendab, et minimaalsuse põhimõtte rakendamine ei toimu enam selgelt etteulatuva kontrollina, vaid pigem tagantjärele antava hinnanguna selle kohta, kas andmete kasutamine oli põhjendatud.

Eesti õiguskirjanduses on samuti arutatud, et GDPR-i põhimõtete kohaldamine tehisintellekti süsteemidele on ebaselge ning võib tekitada märkimisväärset õiguselgusetust, mis omakorda mõjutab nii innovatsiooni kui ka õiguste kaitset.³¹ Samas on praktikas juba märgitud, et just minimaalsuse ja eesmärgipärasuse põhimõtete järgimine on AI kontekstis eriti problemaatiline, kuna süsteemid võivad töödelda andmeid mitmel erineval ja dünaamilisel eesmärgil.³²

Tõepoolest on olukord käesolevaks hetkeks muutunud, sest kui varem määras eesmärk, kui palju andmeid tohib koguda, siis LLM-ide puhul on see pigem vastupidi ja olemasolev andmehulk hakkab määrama, milleks neid kasutada saab. See ei ole üksnes tehniline eripära, vaid mõjutab otseselt ka seda, kuidas õigust kohaldada. Kui eesmärk ei ole piisavalt selgelt määratletud, siis ei ole võimalik hinnata ka seda, milline andmehulk on vajalik ning minimaalsuse põhimõtte praktiline sisu muutub seetõttu määratlematuks.

²⁸ Sartor, G., & Lagioia, F., *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, European Parliament, EPRS, 2020, <https://doi.org/10.2861/293>.

²⁹ <https://smarti.live/1709/eu-mps-ai-ban-work-devices-privacy>

³⁰ Biega, A. J., & Finck, M., *Reviving purpose limitation and data minimisation in data-driven systems, Technology and Regulation* 2021, 44–61, <https://doi.org/10.26116/techreg.2021.004>

³¹ Kaimar Karu, „Tehisintellekti keerukad küsimused“, *Juridica* 2021/1, lk 44, kättesaadav: <https://ojs.utlib.ee/index.php/juridica/article/view/18277>

³² NJORD Estonia, *Is it impossible to be GDPR-compliant in case of AI?*, NJORD Law Firm, 2019, kättesaadav: <https://www.njordlaw.com/print/pdf/node/879>

Tuleb ka arvestada, et tehisintellekti süsteemide puhul ei ole eesmärgid staatilised, vaid ajas muutuvad. Mudelite arendamine, treenimine ja hilisem kasutamine võivad hõlmata erinevaid ja üksteisest sõltuvaid eesmärke, mida ei ole võimalik täielikult ette näha. Selline dünaamilisus ei sobitu hästi GDPR-i eeldusega, et töötlemise eesmärgid tuleb määratleda enne andmete kogumist. Nagu ka empiirilised uuringud näitavad, puudub seni ühtne ja teostatav (nõo operatsionaliseeritav) arusaam sellest, kuidas minimaalsuse põhimõtet keerukates masinõppesüsteemides tegelikult rakendada.³³

Eeltoodust ei tulene, et nimetatud põhimõtted oleksid tehisintellekti kontekstis kohaldamatud, pigem osutab see vajadusele nende sisuliseks ümbertõlgendamiseks. Ühe võimaliku lahendusena on pakutud, et minimaalsuse põhimõtet tuleks tõlgendada mitte rangelt kvantitatiivse piiranguna, vaid funktsionaalse nõudena, mis keskendub sellele, kas konkreetne andmehulk on süsteemi toimimiseks põhjendatud ja proportsionaalne. Näiteks on välja töötatud ka tehnilisi lähenemisi, mis võimaldavad määrata nn optimaalse andmemahu taseme, mille ületamisel ei ole täiendaval andmemahul enam märkimisväärset mõju mudeli jõudlusele.³⁴ Sellised lahendused osutavad, et minimaalsuse põhimõte võib olla rakendatav ka AI-kontekstis, kuid eeldab selle tihedamat ja selgemat sidumist tehniliste kriteeriumidega.

Teiseks võiks kaaluda eesmärgipärasuse põhimõtte erinevat kohaldamist arendus- ja kasutusfaasis. Kui LLM-ide arendamisel võib olla vältimatu lähtuda laiemast ja üldisemast eesmärgist, siis nende kasutamisel konkreetsetes teenustes peaks eesmärk olema selgelt ja kitsalt määratletud. Selline kaheastmeline lähenemine võimaldaks ühelt poolt säilitada innovatsiooni arendusetapis, kuid teiselt poolt tagada rangem andmekaitse siis, kui süsteemi reaalselt kasutatakse.

Kolmandaks on põhjendatud nihutada regulatiivset fookust minimaalsuse formaalselt järgimiselt riskipõhisele hindamisele. Kui andmete täielik minimeerimine ei ole realistlik, muutub keskseks küsimuseks see, milliseid riske andmete töötlemine kaasa toob ning kuidas

³³ Biega, A. J., Potash, P., Daumé III, H., Diaz, F., & Finck, M. (2020). *Operationalizing the legal principle of data minimization for personalization*. arXiv. <https://arxiv.org/abs/2005.13718>

³⁴ Shanmugam, D., Shabarian, S., Diaz, F., Finck, M., & Biega, A. J. (2022). *Learning to limit data collection via scaling laws: A computational interpretation for the legal principle of data minimization*. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. <https://doi.org/10.1145/3531146.3533148>

neid maandada. Nimetatud lähenemine on kooskõlas GDPR-i üldise riskipõhise loogikaga, kuid nõuab selle senisest süsteemsemat rakendamist tehisintellekti kontekstis.

2.3. LLM-ide treeningpraktikad ja andmekogumise läbipaistmatus

LLM-ide arendamine põhineb ulatuslikel ja mitmekesistel andmekogumitel, mille kogumise ja töötlemise praktikad erinevad oluliselt traditsioonilisest andmetöötlustest. Kui GDPR eeldab, et isikuandmete kogumine on eesmärgipärane, läbipaistev ja piiratud, siis LLM-ide arenduses domineerib vastupidine loogika – andmeid kogutakse võimalikult suures mahus, sageli automatiseeritud meetodite abil ning ilma täieliku kontrollita nende päritolu, kvaliteedi või sisulise koosseisu üle. Selline lähenemine ei tekita üksnes praktilisi raskusi, vaid seab kahtluse alla GDPR-i kesksete põhimõtete, eelkõige läbipaistvuse ja minimaalsuse sisulise kohaldatavuse.

LLM-ide treeningandmete kogumisel kasutatakse laialdaselt avalikult kättesaadavat veebisisu, sealhulgas automatiseeritud andmekogumise meetodeid, mida nimetatakse veebikraapimiseks (*web scraping*). Veebikraapimine seisneb andmete süstemaatilises ja automatiseeritud kogumises veebilehtedelt spetsiaalsete tarkvaralahenduste abil, võimaldades ulatuslike andmemahtude koondamist, kuid tõstatades seejuures ka mitmeid andmekaitse- ja autoriõiguslikke küsimusi. Näiteks Common Crawl tüüpi andmekogumid sisaldavad miljardeid veebilehti, mille hulgas leidub paratamatult ka isikuandmeid. Kuigi neid andmeid käsitletakse sageli kui avalikult kättesaadavaid, ei tähenda see nende piiramatut töödeldavust andmekaitseõiguslikus tähenduses. Õiguslikult ei ole määrav üksnes andmete kättesaadavus, vaid ka nende töötlemise eesmärk ja kontekst. Veebist kogutud andmestikud sisaldavad sageli kontrollimata ja problemaatilist sisu, mille päritolu ja korrektsus ei ole täielikult jälgitav. Lisaks tõstab selliste andmete kasutamine küsimuse nende riskasutatavusest, kuna erinevatest allikatest pärinevate andmete kombineerimine võib võimaldada üksikisikute kohta järelduste tegemist viisil, mis ei olnud algselt ette nähtud.³⁵ Selle üheks näiteks on isikute profileerimine ka üksnes avalikult kättesaadavate andmete alusel, mis võib suurendada eraelu puutumatuse riive riski.

³⁵ Dodge, J. jt, *Documenting large webtext corpora: A case study on the Colossal Clean Crawled Corpus*, arXiv:2104.08758 (2021), kättesaadav: <https://arxiv.org/abs/2104.08758>

Siinkohal ilmneb läbipaistvuse põhimõtte üks peamisi probleeme. GDPR eeldab, et andmesubjekt on teadlik oma andmete töötlemisest ning tal on võimalik oma õigusi teostada. LLM-ide kontekstis on see aga praktiliselt teostamatu. Andmesubjektidel puudub reaalne võimalus tuvastada, kas nende andmeid on kasutatud mudelite treenimiseks, ning veelgi enam – puudub mehhanism, mille kaudu nad saaksid nõuda nende andmete eemaldamist juba treenitud mudelist. Samuti ei ole arendajatel endil sageli täielikku ülevaadet kasutatavate andmekogumite täpsest sisust, arvestades nende mahtu ja heterogeensust.³⁶ Selline olukord tähendab, et läbipaistmatus ei ole juhuslik kõrvalnähtus, vaid LLM-ide arenduspraktikate süsteemne tagajärg.

Tõsi, et probleem ei seisne ainult selles, et tehnoloogia on keerukas, vaid ka selles, kuidas reeglid päriselus toimivad. Kui läbipaistvust ei ole tegelikult võimalik sisuliselt tagada, jääb see paratamatult pigem formaalsuseks – kasutajatele antakse küll infot, kuid see piirdub sageli üldiste teadete ja privaatsustingimustega, millest on tavakasutajal raske aru saada. Selline näiline läbipaistvus ei anna inimestele tegelikku kontrolli oma andmete üle, kuigi see on GDPR-i üks peamisi eesmärke. Seetõttu võib öelda, et suurte keelemudelite puhul on läbipaistvuse põhimõte vähemalt osaliselt oma tähenduse kaotanud.

Probleem ei piirdu üksnes andmete kogumisega, vaid laieneb ka nende valiku ja töötlemise etappidele. Kuigi arendajad kasutavad erinevaid filtreerimismeetodeid sobimatu sisu eemaldamiseks, ei ole need protsessid täielikult läbipaistvad ega ammendavad. Uuringud on näidanud, et ka pärast andmete puhastamist võivad treeningandmestikesse jääda isikuandmed või muidu eksitav teave, mis võib hiljem avalduda mudeli väljundites.³⁷ Näitena saaks siia tuua olukorra, kui tehisintellekti mudel on treenitud suure hulga kontrollimata andmete peal ja sinna on sattunud ka kellegi vana foorumipostitus, kus ta jagas oma telefoninumbrit koos nimega. Tavaliselt mudel seda infot niisama ei tea, aga kui keegi küsib piisavalt täpselt (nt mis on selle inimese kontakt?), võib mudel juhuslikult selle numbri välja pakkuda, sest see on tal treeningust meelde jäänud. Probleem on selles, et inimene ise ei tea, et tema andmed on sellises süsteemis sees, ega saa seda kontrollida või kustutada. Just seetõttu ei piisa ainult sellest, et kogutakse

³⁶ Dodge, J. jt, *Documenting large webtext corpora: A case study on the Colossal Clean Crawled Corpus*, arXiv:2104.08758 (2021), kättesaadav: <https://arxiv.org/abs/2104.08758>

³⁷ Birhane, A., Prabhu, V. U., & Kahembwe, E., *Multimodal datasets: Misogyny, pornography, and malignant stereotypes*, arXiv:2110.01963 (2021), kättesaadav: <https://arxiv.org/abs/2110.01963>

vajalikud andmed – oluline on ka hinnata, kas ja kuidas need liigsed andmed võivad hiljem ootamatult välja lekkida.

Täiendava keerukuse toob kaasa LLM-ide arenduse hajutatud iseloom. Andmekogumite loomine, töötlemine ja kasutamine toimub sageli mitme osapoole koostöös, kus andmete päritolu ja töötlemise tingimused võivad aja jooksul muutuda. Selline astmeline andmetöötlus raskendab nii vastutuse määratlemist kui ka läbipaistvuse tagamist, kuna puudub ühtne kontrollpunkt, mille kaudu oleks võimalik kogu protsessi hinnata. Eesti õiguskirjanduses on samuti leitud, et selline hajus andmetöötlus võib viia olukorrani, kus GDPR-i nõuete järgimine muutub sisuliselt raskesti kontrollitavaks.³⁸

Tehisintellekti, eriti LLM-ide kontekstis eeldab läbipaistvuse põhimõtte tõhus rakendamine fookuse nihutamist üksikisikute teavitamiselt süsteemi tasandile. Kui andmeid kogutakse ulatuslikult ja mitmekesistest allikatest, ei ole individuaalne teavitamine sageli praktiliselt teostatav ega taga andmesubjekti tegelikku kontrolli oma andmete üle.³⁹ Seetõttu tuleb läbipaistvust käsitada eelkõige süsteemse omadusena, mis väljendub andmekogumite, arendusprotsesside ja riskide struktureeritud dokumenteerimises ning nende sõltumatus kontrollitavuses. Selline lähenemine võimaldab hinnata andmetöötluste õiguspärasust ja mõju tervikuna, mitte üksnes üksikute teavitustoimingute kaudu.

Eeltoodust tulenevalt on põhjendatud küsida, kas läbipaistvuse põhimõtet on võimalik LLM-ide kontekstis rakendada samal viisil nagu traditsioonilises andmetöötlustes. Minu hinnangul eeldab see põhimõtte sisulist ümbertõlgendamist. Ühe võimaliku lähenemisena võiks rõhutada individuaalse teavitamise asemel süsteemset läbipaistvust, mis hõlmab andmekogumite dokumenteerimist, arendusprotsesside kirjeldamist ning riskide avalikustamist. Näiteks nn andmestike kirjelduslehtede lähenemine (*datasheets for datasets*) võimaldab andmestike päritolu ja koostist süsteemsemalt kirjeldada, kuigi see ei lahenda täielikult andmesubjektide individuaalsete õiguste teostamisega seotud probleeme. Nimetatud lähenemine seisneb selles, et andmestike kohta koostatakse standardiseeritud kirjeldused, milles selgitatakse nende päritolu, koostist ning võimalikke riske.⁴⁰

³⁸ Tupay, P. K., Ebers, M., Juksaar, J., & Kohv, K. (2021). *Is European data protection toxic for innovative AI? An Estonian perspective*. Juridica International, 30, 141–153.

³⁹ Gebru, T., Morgenstern, J., Vecchione, B., Wortman Vaughan, J., Wallach, H., Daumé III, H., & Crawford, K. (2018). *Datasheets for datasets*. arXiv:1803.09010. <https://arxiv.org/abs/1803.09010>

⁴⁰ *Ibid.*

Samuti võiks kaaluda regulatiivsete nõuete täpsustamist selliselt, et need arvestaksid LLM-ide arenduse eripäradega. Kui täielik läbipaistvus ei ole tehniliselt saavutatav, tuleks keskenduda sellele, milline läbipaistvuse tase on piisav, et tagada andmesubjektide õiguste sisuline kaitse. See võib tähendada suurema rõhu panemist järelevalveasutuste rollile, standardiseeritud dokumentatsiooninõuetele ning sõltumatutele audititele.

2.4. Proportsionaalsuse hindamine ja järelevalvepraktika

Proportsionaalsuse põhimõte kujutab endast Euroopa Liidu õiguse alustala, mille kaudu hinnatakse põhiõiguste riivete lubatavust ning õigustatakse avaliku ja eraõigusliku sekkumise ulatust. Kuigi GDPR ei sätesta proportsionaalsust eraldi põhimõttena, on see sisuliselt integreeritud mitmetesse andmekaitseõiguse normidesse, eelkõige andmete minimaalsuse, eesmärgipärasuse ja õiguspärasuse nõuetesse. Proportsionaalsuse test koosneb klassikaliselt kolmest astmest – sobivus, vajalikkus ja mõõdukus –, mille kaudu hinnatakse, kas konkreetne andmetöötlus on õigustatud. Tehisintellekti, eriti suurte keelemudelite kontekstis muutub aga selle testi rakendamine märkimisväärselt keerukamaks, kuna puudub selge ja kontrollitav seos kasutatavate andmete ja saavutatud tulemuste vahel.⁴¹

Sobivuse kriteeriumi kohaldamisel tekib küsimus, kas ulatuslik andmekogumine on üldse sobiv vahend seatud eesmärgi saavutamiseks. LLM-ide puhul võib esmapilgul väita, et mida suurem on andmekogum, seda parem on mudeli võimekus. Kuid selline eeldus ei ole pädev, et õigustada piiramatut andmekogumist. Kui sobivust käsitatakse liiga üldisel tasandil, muutub see sisutühjaks, mis ei täida enam oma piiravat funktsiooni.

Vajalikkuse kriteerium tähendab, et andmete töötlemine peab olema eesmärgi saavutamiseks hädavajalik ning sama tulemust ei tohi olla võimalik saavutada leebemate vahenditega, näiteks ilma isikuandmeid kasutamata või ainult mitteisikuliste andmete abil. LLM-ide kontekstis on selline alternatiivide hindamine aga keeruline, kuna mudelite arendamine tugineb sageli ulatuslikele ja mitmekesistele andmetikele, mistõttu ei ole selge, millises ulatuses oleks võimalik eesmärgi saavutada ilma isikuandmete kasutamiseta. Masinõppesüsteemide puhul ei ole üksikute andmeelementide mõju lõpptulemusele lineaarne ega läbipaistev, mistõttu

⁴¹ European Data Protection Board, *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models* (17 December 2024).

vajalikkuse hindamine muutub suuresti spekulatiivseks.⁴² Sellest saab järeldada, et vajalikkuse test ei saa toimida klassikalises tähenduses ning võib viia olukorrani, kus praktiliselt igasugune andmekogumine loetakse põhjendatuks.

Mõõdukuse ehk proportsionaalsuse kitsamas tähenduses tähendab andmetöötlustest saadava kasu ja sellega kaasnevate riskide kaalumist. LLM-ide puhul on selline kaalumine aga eriti keeruline, kuna nende mõju ei ole piiratud üksikjuhtumitega, vaid võib avalduda laialdaselt ja ettearvamatult. Lisaks ei ole selliste riskide realiseerumine alati prognoositav, mistõttu on keeruline anda lõplikku hinnangut andmetöötluste mõõdukusele. Näiteks võivad mudelid genereerida ebaõigeid või tundlikke andmeid, mis võivad kahjustada üksikisikute õigusi, kuid selliste riskide realiseerumine ei ole alati prognoositav. Seetõttu muutub ka mõõdukuse hindamine ebamääraseks ning sõltub suuresti hinnangulistest kaalutlustest.

Sama probleem tuleb välja ka järelevalves. Seni on GDPR-i rakendamine tehisintellekti puhul olnud pigem üksikjuhtumite põhine ja tagantjärele reageeriv. See näitab, et järelevalve ei ole veel täielikult tehnoloogia arenguga sammu pidanud.

Eestis on asutud seisukohale, et tehisintellekti regulatsioon eeldab järelevalveasutustelt märkimisväärset tehnilist kompetentsi, ilma milleta ei ole võimalik hinnata süsteemide tegelikku toimimist ega riske.⁴³ Autorid leiavad, et kuigi GDPR näeb ette andmekaitseasutuste keskse rolli isikuandmete töötlemise järelevalves, tekitab tehisintellekti rakendamine praktikas olulist õigusselguse puudujääki, kuna normide kohaldamine keerukatele ja dünaamilistele AI-süsteemidele ei ole üheselt määratletav. Ebaühtlane ja ettearvamatu järelevalvepraktika liikmesriikides võib suurendada õiguslikku ebakindlust ning seeläbi pärssida innovaatiliste AI-lahenduste kasutuselevõttu. Samas rõhutatakse, et järelevalve roll ei seisne innovatsiooni takistamises, vaid põhiõiguste kaitse tagamises olukorras, kus tehnoloogiline areng ületab kehtiva regulatsiooni selguse.

⁴² Wachter, S., Mittelstadt, B., & Floridi, L., *Why a right to explanation of automated decision-making does not exist in the GDPR*, *International Data Privacy Law* 7(2), 76–99 (2017), <https://doi.org/10.1093/idpl/ix005>

⁴³ Tupay, P. K., Ebers, M., Juksaar, J., & Kohv, K., *Is European data protection toxic for innovative AI? An Estonian perspective*, *Juridica International* 30 (2021), 141–153, <https://doi.org/10.12697/JI.2021.30.12>, kättesaadav: https://www.juridicainternational.eu/article_full.php?uri=2021_30_is_european_data_protection_toxic_for_innovative_ai_an_estonian_perspective.

Seega, kui järelevalveasutus ei suuda mõista, kuidas mudel andmeid töötleb või millised on selle võimalikud väljundid, muutub ka proportsionaalsuse hindamine sisuliselt formaalseks.

Eeltoodud probleemide valguses on põhjendatud otsida alternatiivseid lähenemisi proportsionaalsuse põhimõtte rakendamiseks. Seejuures ei välista miski proportsionaalsuse põhimõtte kohaldamist süsteemsel tasandil, käsitledes andmetöötlust tervikuna ning diferentseerides seda nii andmeliikide kui ka -mahu alusel. Selline lähenemine eeldab, et hinnatakse mitte üksnes üksikute andmeelementide vajalikkust, vaid kogu andmetöötluse õigustatust ning selle mõju andmesubjektide õigustele. Teiseks võiks tugevdada ex ante (e.k) kontrollimehhanisme, eelkõige andmekaitsealase mõjuhinnangu (DPIA) rolli. Kui proportsionaalsuse hindamine ei ole võimalik üksnes ex post (e.p) analüüsi kaudu, tuleb suuremat rõhku panna riskide ennetavale hindamisele. See eeldab aga, et DPIA-d ei jääks formaalseks dokumendiks, vaid sisaldaksid sisulist analüüsi ja tehnilist põhjendust miks ja kuidas andmeid kasutatakse. Kolmandaks on vajalik arendada järelevalvepraktikat, mis oleks kohandatud tehisintellekti eripäradele. See võiks hõlmata standardiseeritud hindamiskriteeriume, dokumenteerimist, sõltumatuid auditeid ning tihedamat koostööd tehniliste ekspertidega. Samas tuleb kriitiliselt märkida, et sellised meetmed ei kõrvalda täielikult proportsionaalsuse hindamisega seotud ebakindlust, vaid aitavad seda üksnes vähendada.

Samal ajal on Euroopa Andmekaitse nõukogu (European Data Protection Board, EDPB) rõhutanud, et tehisintellekti süsteemide hindamisel ei piisa üksnes formaalsest vastavusest GDPR-i nõuetele, vaid vajalik on sisuline analüüs selle kohta, milline on süsteemi tegelik mõju andmesubjektide õigustele.⁴⁴ EDPB seisukohtadest ilmneb, et proportsionaalsuse hindamine peaks hõlmama kogu andmetöötluse elutsükli, sealhulgas andmete kogumist, mudeli treenimist ja selle rakendamist. Seejuures tuleb iga etappi hinnata eraldi, arvestades selle eripära ja sellega seotud riske, mis ei välista andmetöötluse terviklikku hindamist kogu elutsükli ulatuses. Samuti on juhitud tähelepanu sellele, et suurte keelemudelite puhul võib riskide hindamine olla keeruline just nende läbipaistmatuse ja ettearvamatus tõttu.

Täiendavalt on Euroopa Andmekaitseinspektor (European Data Protection Supervisor, EDPS) rõhutanud vajadust tugevdada ennetavaid kontrollimehhanisme, sealhulgas DPIA rolli ning

⁴⁴ European Data Protection Board, *Statement on generative artificial intelligence and data protection* (2023), kättesaadav: https://www.edpb.europa.eu/news/news/2023/edpb-statement-generative-ai_en.

sõltumatut järelevalvet. EDPS on toonud esile, et traditsiooniline proportsionaalsuse hindamine ei pruugi olla piisav olukordades, kus tehnoloogilised süsteemid arenevad kiiremini kui regulatiivsed raamistikud. Seetõttu on soovitatud liikuda suurema standardiseerituse ja tehnilise kontrollitavuse suunas, et tagada õigusnormide tegelik rakendatavus.

Euroopa Liidu institutsioonide seisukohad seega kinnitavad, et proportsionaalsuse põhimõte ei ole tehisintellekti puhul kadunud, vaid selle rakendamine on muutunud keerukamaks ja mitmetasandiliseks. Samas tekib ka vastuolu, sest kuigi räägitakse paindlikust ja riskipõhisest lähenemisest, ei ole veel selgeid reegleid, mille järgi hinnata, kas LLM-ide andmekasutus on proportsionaalne. See näitab, et regulatsiooni tuleb edasi arendada, et õiguslikud põhimõtted ja tehnilised hindamismeetodid paremini kokku viia.

2.5. Tehnilised ja korralduslikud meetmed andmete minimaalsuse ja läbipaistvuse toetamiseks

Tehniliste ja korralduslike meetmete kaudu realiseeritakse abstraktsed andmekaitsepõhimõtted, sealhulgas andmete minimaalsus ja läbipaistvus. GDPR artikkel 5 ei piirdu üksnes põhimõtete deklaratiivse sätestamisega, vaid eeldab nende sisulist rakendamist läbi konkreetsete mehhanismide ja praktikate.⁴⁵ Tehisintellekti, eelkõige LLM-ide kontekstis omandavad need meetmed aga erilise tähenduse, kuna traditsioonilised andmekaitsemehhanismid ei ole sageli piisavad, et tagada põhimõtete tegelik järgimine.

Esiteks tuleb esile tõsta nn lähteandmekaitse ja vaikimisi andmekaitse põhimõtteid (*privacy by design* ja *privacy by default*), mille kohaselt tuleb andmekaitse integreerida süsteemi juba selle kavandamise ja arendamise etapis.⁴⁶ See eeldab, et andmekaitse ei ole tehisintellekti süsteemide puhul pelgalt järelevalve- või vastavuskontrolli objekt, vaid süsteemi arhitektuuri ja toimimise lahutamatu osa. Isikuandmete kaitse tuleb integreerida tehnoloogiate ja äriprotsesside kavandamisse juba varases disainifaasis, tagades vaikimisi privaatsuse kogu andmetöötluse elutsükli ulatuses. Selline lahendus nihutab fookuse tagantjärele sekkumiselt ennetavale riskide maandamisele ning võimaldab siduda andmete minimaalsuse ja vastutuse põhimõtted tehniliste

⁴⁵ Regulation (EU) 2016/679 (GDPR), art 5 – *Principles relating to processing of personal data*, veebist kättesaadav: <https://gdpr-info.eu/art-5-gdpr>

⁴⁶ Cavoukian, A., *Privacy by design: The 7 foundational principles*, Information and Privacy Commissioner of Ontario, 2011, kättesaadav: <https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf>

ja korralduslike lahendustega, mis toetavad nii põhiõiguste kaitset kui ka süsteemide usaldusväärset toimimist.

Andmete minimaalsuse põhimõtte järgimine ei pea piirduma abstraktse normatiivse nõudega, vaid seda on võimalik toetada konkreetsete tehniliste lahendustega, mis aitavad arendajatel süsteemselt hinnata, milline andmete osa on konkreetse süsteemi eesmärkide saavutamiseks tegelikult vajalik.⁴⁷ Selliste lähenemiste puhul ei lähtuta enam eeldusest, et andmeid tuleks koguda võimalikult palju, vaid tehakse teadlik ja põhjendatud valik nii andmeliikide kui ka andmemahu osas, lähtudes nende vajalikkusest ja proportsionaalsusest konkreetse eesmärgi suhtes.

Teiseks on arendatud mitmeid tehnilisi minimaalsuse tööriistu, mis püüavad teostada õiguslikku põhimõtet masinõppe kontekstis. Näiteks on pakutud mudeleid, mis võimaldavad hinnata andmete kogumise ja mudeli jõudluse vahelist seost ning määrata nn optimaalse andmemahu, mille ületamine ei anna enam olulist lisaväärtust. Samuti on välja töötatud meetodeid, mis võimaldavad vähendada sisendandmete hulka või üldistada neid ilma mudeli täpsust oluliselt kahjustamata.⁴⁸ Taolised võimalikud meetodid ilmestavad, et minimaalsuse põhimõtte ei ole tehniliselt võimatu, vaid pigem alarakendatud.

Siiski tuleb kriitiliselt märkida, et sellised tehnilised lahendused ei lahenda minimaalsuse probleemi tervikuna. Andmepõhiste süsteemide puhul on tegemist olukorraga, kus andmete kogumine on tihedalt seotud süsteemi toimimisloogikaga ning minimaalsuse rakendamine võib tekitada ootamatuid kompromisse mudeli täpsuse ja õigsuse vahel.⁴⁹

Nõustun, et probleem tekib olukorras, kus andmete vähendamine minimaalsuse põhimõtte järgimiseks mõjutab süsteemi toimimist viisil, mida ei ole võimalik hinnata üksnes privaatsuse seisukohast. Väiksem andmemahut võib vähendada mudeli täpsust või avaldada ebaproportsionaalset mõju teatud isikurühmadele, mistõttu võivad tulemused muutuda ebaühtlasemaks või isegi kallutatuks. Seetõttu ei saa andmete minimaalsust käsitada pelgalt tehnilise optimeerimisülesandena, vaid see eeldab erinevate õigushüvede – eelkõige eraelu

⁴⁷ Iwaya, L. H., Babar, M. A., & Rashid, A. (2022). *Privacy engineering in the wild: Understanding the practitioners' mindset, organisational culture, and current practices*. IEEE Security & Privacy / arXiv preprint. <https://arxiv.org/abs/2211.08916>

⁴⁸ Goldstein, A. J., *Data minimization for GDPR compliance in machine learning models*, arXiv:2008.04113 (2020), veebist kättesaadav: <https://arxiv.org/abs/2008.04113>

⁴⁹ Biega, A. J., & Finck, M., *Reviving purpose limitation and data minimisation in data-driven systems*, arXiv:2101.06203 (2021), veebist kättesaadav: <https://arxiv.org/abs/2101.06203>

puutumatus, otsuste täpsuse ja võrdse kohtlemise – tasakaalustatud kaalumist. Näiteks võib minimaalsuse põhimõtte rakendamine tähendada, et mudeli treenimisel jäetakse välja teatud muutujad, mida ei peeta esmapilgul vajalikuks. Kuigi see vähendab isikuandmete töötlemise ulatust, võib see samal ajal piirata süsteemi võimet eristada erinevaid olukordi või isikurühmi, mille tagajärjel võib langeda otsuste täpsus või suurendada ebavõrdse kohtlemise risk. Sellisel juhul ei avaldu vastuolu abstraktselt privaatsuse ja efektiivsuse vahel, vaid konkreetsetes kompromissis andmete minimaalsuse ning süsteemi usaldusväärsuse ja õiguspärase toimimise vahel. Sellest tulenevalt ei ole üksnes tehnilised meetmed piisavad minimaalsuse põhimõtte täielikuks tagamiseks, vaid neid tuleb käsitada vahenditena, mis aitavad riske maandada, kuid ei kõrvalda vajadust sisulise õigusliku kaalumise järele.

Läbipaistvuse suurendamiseks andmepõhistes süsteemides on välja töötatud mitmesuguseid dokumenteerimis- ja selgitusmehhanisme, mille eesmärk on muuta nii andmekogumite kui ka algoritmiliste mudelite kasutus paremini mõistetavaks. Näiteks on andmestike kirjelduslehtede lähenemine (*datasheets for datasets*) suunatud andmekogumite päritolu, koostise, kavandatud kasutusviiside ning võimalike riskide süstemaatilisele kirjeldamisele, toetades seeläbi vastutustundlikumat andmekäitlust.⁵⁰ On ka arendatud nn selgitatava tehisintellekti (*explainable AI, XAI*) meetodeid, mis püüavad avada mudelite otsustusloogikat ja muuta automatiseeritud otsused arusaadavamaks nii arendajatele kui ka järelevalveasutustele.⁵¹ Samas täidavad sellised läbipaistvusmehhanismid sageli eeskätt formaalse vastavuse eesmärki ega pruugi tagada andmesubjektidele sisulist arusaamist süsteemide tegelikust toimimisest. Seetõttu ei ole läbipaistvuse probleem üksnes tehniline, vaid hõlmab ka kommunikatiivseid ja institutsionaalseid aspekte, sealhulgas seda, kellele ning millisel detailsus- ja arusaadavuse tasemel selgitusi antakse, arvestades erinevate adressaatide (nt andmesubjektid, järelevalveasutused) vajadusi.⁵²

Eesti õiguskorra kontekstis ei tohiks läbipaistvust andmepõhiste ja tehisintellektil põhinevate süsteemide puhul käsitada üksnes tehnilise dokumenteerimise või mudeli selgitamise küsimusena, vaid osana avaliku võimu vastutusmehhanismidest. Arvestades, et paljud Eestis kasutatavad digilahendused toimivad avaliku sektori teenuste või otsuste ettevalmistamise

⁵⁰ Gebru, T., Morgenstern, J., Vecchione, B., Wortman Vaughan, J., Wallach, H., Daumé III, H., & Crawford, K. (2018). *Datasheets for datasets*. arXiv:1803.09010. <https://arxiv.org/abs/1803.09010>

⁵¹ Mittelstadt, B., Russell, C., & Wachter, S. (2019). *Explaining explanations in AI*. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAT*). <https://arxiv.org/abs/1811.01439>

⁵² Wachter, S., Mittelstadt, B., & Floridi, L. (2017). *Why a right to explanation of automated decision-making does not exist in the GDPR*. *International Data Privacy Law*, 7(2), 76–99. <https://doi.org/10.1093/idpl/ix005>

tasandil, peaks läbipaistvuse tagamine olema suunatud eeskätt sellele, et nii andmesubjektidel kui ka järelevalveasutustel oleks võimalik mõista süsteemi rolli otsustusprotsessis ja hinnata selle mõju õigustele. See eeldab lisaks tehnilistele selgitus- ja dokumenteerimisvahenditele ka institutsionaalseid valikuid, näiteks selget vastutuse jaotust arendajate, teenuseosutajate ja haldusorganite vahel ning arusaadavat kommunikatsiooni selle kohta, millisel määral automatiseeritud süsteeme tegelikult otsustamises kasutatakse. Kirjeldatud lähenemine võimaldaks vältida olukorda, kus läbipaistvus taandub formaalseks nõude täitmiseks, kuid ei toeta sisuliselt õiguspärast ja kontrollitavat haldustegevust.

Korralduslikest meetmetest on olulised andmekaitsealased mõjuhinnangud, sisekontrollisüsteemid ja auditid, mille eesmärk on hinnata ja dokumenteerida andmetöötlusega seotud riske. Tehisintellekti kontekstis on aga probleemiks see, et sellised mehhanismid võivad praktikas taanduda formaalseks vastavuseks, kui nende aluseks olev tehniline riskihinnang ja süsteemne analüüs on ebapiisavad. Nagu on leitud ja näidatud ka AI regulatsiooni käsitlevates uuringutes, eeldab tõhus järelevalve interdistsiplinaarset lähenemist, kus õiguslikud nõuded on sisuliselt seotud tehnilise arusaamaga AI-süsteemide toimimisest.⁵³

Paraku ei ole probleem niivõrd selles, et tehnilisi ja korralduslikke meetmeid ei oleks olemas, vaid pigem selles, kuidas neid tegelikult kasutatakse. Sageli rakendatakse neid alles hiljem nn kaitsekihina, kui süsteemi põhiloogika –võimalikult ulatuslik andmete kogumine ja piiratud läbipaistvus – on juba paika pandud. Fikseeritud olukorras on süsteemi ülesehitus juba selline, et selle toimimist ei ole enam lihtne sisuliselt kontrollida ega hinnata. Eeltoodust tulenevalt leian, et nii DPIA-de kui ka laiema järelevalve toimimiseks ei piisa pelgalt formaalsest kontrollist. Vaja on sisulist analüüsi ja tehnilist põhjendust, mis võimaldaks päriselt hinnata, kuidas süsteem toimib ja millist mõju see inimestele avaldab.

Euroopa Liidu tasandil on tehniliste ja korralduslike meetmete rolli tehisintellekti reguleerimisel kirjeldatud eeskätt riskipõhise regulatiivse lähenemise kaudu. Euroopa Komisjon on tehisintellekti regulatsiooni kujundamisel lähtunud eeldusest, et kõrge riskiga

⁵³ Hohmann, B., & Kollár, G., *Reflections on the data protection compliance of AI systems under the EU AI Act*, *Cogent Social Sciences* 11(1) (2025), art 2560654, DOI: 10.1080/23311886.2025.2560654, veebist kättesaadav: <https://www.tandfonline.com/doi/full/10.1080/23311886.2025.2560654>

AI-süsteemid peavad vastama rangematele nõuetele, mis hõlmavad muu hulgas andmete kvaliteeti, ulatuslikku dokumenteerimist, läbipaistvust ning inimjärelevalve tagamist.⁵⁴

Samas on Euroopa Andmekaitsekoostöögrupp selgitanud, et tehnilised meetmed ei saa asendada vastutava töötleja õiguslikku vastutust, vaid peavad toimima seda toetava mehhanismina, mille tõhusus sõltub selgetest organisatsioonilistest protsessidest, sealhulgas dokumenteerimisest, riskihindamisest ja järelevalvest.⁵⁵ Täiendavalt on Euroopa Andmekaitse Inspektor toonud esile vajaduse tugevdada AI-süsteemide kontrollitavust ja standardiseerimist, et tagada nende tegelik kontrollitavus ja vältida pelgalt formaalset vastavust.⁵⁶ Kontrollitavuse all tuleb tehisintellekti kontekstis mõista süsteemi võimet olla sõltumatult ja tõenduspõhiselt kontrollitav kogu selle elutsükli vältel, võimaldades hinnata nii tehniliste kui ka korralduslike meetmete tegelikku vastavust andmekaitse nõuetele. Standardiseerimise kaudu on võimalik muuta andmete minimaalsuse ja läbipaistvuse põhimõtted sisuliselt rakendatavaks, kehtestades ühtsed nõuded andmekogumite kvaliteedile, dokumenteerimisele ja kontrollitavusele, mille alusel saab tehisintellekti süsteemide vastavust hinnata võrreldaval ja järelevalvele alluval viisil. Seetõttu tehnilised ja korralduslikud meetmed ei ole enam üksnes soovituslikud head tavad, vaid kujunevad üha selgemalt normatiivseks eeltingimuseks tehisintellekti õiguspärasele kasutamisele.

Eesti digiriigi kiire areng näitab, et tehnoloogiliste lahenduste kasutuselevõtt avalikus sektoris on sageli olnud kantud tõhususe ja automatiseerimise loogikast, samas kui andmekaitse kaalutlused on jäänud pigem järelevalve ja tagantjärele kontrolli tasandile. Tehisintellekti rakendamise puhul ei ole selline lähenemine enam piisav, kuna süsteemide keerukus ja andmetöötluse ulatus muudavad hilisema sekkumise nii tehniliselt kui ka õiguslikult problemaatiliseks. Seetõttu tuleks andmekaitse integreerimist Eestis käsitleda mitte üksnes formaalse vastavusnõudena, vaid strateegilise disainipõhimõttena, mis suunab juba varases arendusfaasis otsuseid andmete valiku, kasutusviiside ja riskide maandamise kohta.⁵⁷ Selline lähenemine võimaldaks säilitada Eesti digiriigi innovatsioonivõime, vältides samas olukorda,

⁵⁴ European Commission. (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. COM(2021) 206 final. Veebist kättesaadav: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>

⁵⁵ European Data Protection Board. (2023). *Statement on generative artificial intelligence and data protection*. Veebist kättesaadav: https://www.edpb.europa.eu/news/news/2023/edpb-statement-generative-ai_en

⁵⁶ European Data Protection Supervisor. (2020). *Opinion 4/2020 on the European Commission's White Paper on Artificial Intelligence – A European approach to excellence and trust*. Veebist kättesaadav: https://www.edps.europa.eu/sites/default/files/publication/20-06-19_opinion_ai_white_paper_en.pdf

⁵⁷ Riigi Infosüsteemi Amet, *Eesti digiriigi arengukava 2030*, kättesaadav: <https://www.ria.ee>

kus põhiõiguste kaitse taandub deklaratiivseks või sõltub üksnes järelevalveasutuste sekkumisest.

Menetluses olevate haldusmenetluse seaduse ja avaliku teabe seaduse muutmisega soovitakse luua selge ja tehnoloogianeutraalne õiguslik raamistik automaatse haldusmenetluse kasutamiseks (*automated administrative decision-making*). Selle all peetakse silmas menetlust, kus haldusakt, muu dokument või toiming tehakse infosüsteemi abil automaatselt, ilma ametniku vahetu sekkumiseta, kasutades selleks näiteks registripäringuid, andmete riskasutust ning etteantud otsustusreeglitel põhinevaid algoritme.⁵⁸ Eelnõu seletuskirjast tulenevalt on muudatuse eesmärk muuta haldusmenetlus kiiremaks, tõhusamaks ja inimestele mugavamaks, eelkõige lihtsamates ja korduvates juhtumites. Samas rõhutatakse, et ka automaatse menetluses peavad olema tagatud haldusmenetluse põhimõtted, sealhulgas õiguspärasus, läbipaistvus ja põhjendatus.

GDPR art 22 ning põhjendus 71 seab piirid täielikult automatiseeritud otsustele (*automated decision-making*) ning nõuab, et isikule oleks tagatud õigus saada selgitusi, nõuda inimlikku sekkumist (*human intervention*) ja otsust vaidlustada. Need nõuded haakuvad otseselt ka käesoleva töö uurimisteemaga, kuna nii automatiseeritud haldusmenetluse kui ka suurte keelemudelite kasutamise puhul kerkivad esile sarnased probleemid seoses andmete minimaalsuse, töötlemise eesmärgipärasuse ning isikuandmete kasutamise ulatuse põhjendatusega. Eelkõige tõusetub küsimus, millises ulatuses on lubatav kasutada ulatuslikke ja mitmekesiseid andmestikke olukorras, kus otsustusprotsess on osaliselt või täielikult automatiseeritud ning selle mõju andmesubjektide õigustele ei ole alati ette täielikult prognoositav.

Siiski tuleb järeldada, et tehnilised ja korralduslikud meetmed on küll vältimatud andmete minimaalsuse ja läbipaistvuse põhimõtete rakendamisel tehisintellekti kontekstis, kuid nende tegelik mõju on olemuslikult piiratud. Probleem ei seisne üksnes konkreetsete meetmete olemasolus või puudumises, vaid eeskätt selles, et need ei pruugi olla võimelised ületama süsteemseid keerukusi ja kontrollimatust, mis tulenevad tehisintellekti arendusloogikast, kus andmete ulatuslik kogumine ja piiratud läbipaistvus on sageli süsteemide toimimise vaikumise eeldus.

⁵⁸ Haldusmenetluse seaduse ja avaliku teabe seaduse muutmise seaduse eelnõu seletuskiri, kättesaadav: <https://eelnou.valitsus.ee>

III PEATÜKK

ANONÜMISEERIMINE JA PSEUDONÜMISEERIMINE TEHISINTELLEKTI RAKENDUSTES

3.1. Anonümiseerimise ja pseudonümiseerimise õiguslik eristus GDPR-is

Anonümiseerimine ja pseudonümiseerimine on GDPR-i süsteemis kaks olulist, kuid praktikas sageli ekslikult samastatud mõistet, mille õiguslik tähendus, regulatiivne funktsioon ja tagajärjed erinevad oluliselt. GDPR artikli 4 lõike 5 kohaselt tähendab pseudonümiseerimine isikuandmete töötlemist viisil, mille tulemusel ei ole andmeid võimalik enam seostada konkreetse isikuga ilma täiendava teabeta, eeldusel et see täiendav teave on hoitud eraldi ja kaitstud asjakohaste tehniliste ja korralduslike meetmetega.⁵⁹ Sellest tulenevalt ei kõrvalda pseudonümiseerimine isikuandmete seost andmesubjektiga, vaid üksnes vähendab tuvastatavuse riski. Kuna seos on põhimõtteliselt taastatav, jäävad pseudonümiseeritud andmed endiselt isikuandmeteks ning alluvad täielikult GDPR-i kohaldamisalale. Anonümiseerimine seevastu tähendab olukorda, kus isiku tuvastamine ei ole enam võimalik ei otseselt ega kaudselt, mistõttu jäävad sellised andmed GDPR-i kohaldamisalast välja. See eristus on normatiivses plaanis selge, kuid praktikas äärmiselt nõudlik.

Euroopa Liidu õiguses on anonümiseerimise standard kujundatud teadlikult kõrgeks, et vältida isikuandmete kaitse sisulist õõnestamist pelgalt formaalsele või näilisele põhjendusele tuginedes. Euroopa Kohtu praktikas, kohtuasjas Breyer v Germany on asutud seisukohale, et isiku tuvastatavuse hindamisel tuleb arvesse võtta kõiki vahendeid, mida on mõistlikult tõenäoline kasutada, sealhulgas kolmandate isikute käsutuses olevaid ressursse ja tehnoloogilisi arenguid.⁶⁰ Kohus leidis antud juhul, et ka IP-aadresse võib käsitada isikuandmetena, kui veebisaidi haldajal on õiguslikud vahendid tuvastada kasutaja, kasutades täiendavat teavet kolmandatelt isikutelt (nt internetiteenuse pakkujalt). Kohtuasja käsitluses ei piirdu tuvastatavus pelgalt sellega, kas andmetöötlejal tehniliselt on võimalik isikut kohe tuvastada. Oluline on ka see, kas tal on reaalne ja õiguspärane ligipääs täiendavale teabele. Kohus rõhutas,

⁵⁹ Regulation (EU) 2016/679, art 4(5) (pseudonümiseerimine) ja põhjendus 26 (anonüümsed andmed), veebist kättesaadav: <https://gdpr-info.eu/art-4-gdpr/> ; <https://gdpr-info.eu/recitals/no-26/>

⁶⁰ CJEU, *Patrick Breyer v Bundesrepublik Deutschland* (C-582/14), ECLI:EU:C:2016:779, p-d 42–46, veebist kättesaadav: <https://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=EN>

et isikuandmete mõistet tuleb tõlgendada laialt, et tagada põhiõiguste kaitse. Nimelt, anonümiseerimine ei ole hinnatav üksnes konkreetse töötaja vaatenurgast ega ainult hetkeolukorra põhjal, vaid nõuab sisulist hinnangut, milles arvestatakse ka tulevikus tekkivaid tuvastamisvõimalusi. Seetõttu ei ole anonümiseerimine pelgalt tehniline protseduur, vaid õiguslik hinnang, mis seob tehnilised meetodid nende tegeliku mõju ja riskiprofiiliga. Eluliselt muudab see anonümiseerimise pigem ideaalseks sihiks kui kergesti saavutatavaks seisundiks.

Anonümiseerimine on oluliselt keerukam, kui definitsioon eeldab. Isegi tugevalt anonümiseeritud andmekogumeid on võimalik taasidentifitseerida, kui neid kombineerida teiste, näiliselt sõltumatute andmeallikatega. Sellised uuringud on toonud esile, et anonümiseerimine ei pruugi olla pöördumatu, eriti andmemahukate ja omavahel ühendatud infosüsteemide kontekstis.⁶¹ See seab kahtluse alla anonümiseerimise kui lõpliku ja universaalse lahenduse andmekaitse probleemidele, eeskätt olukordades, kus andmeid töödeldakse ulatuslikult, pikaajaliselt ja mitme erineva eesmärgi raames. Kuigi anonümiseerimine eeldab otseste identifikaatorite eemaldamist, ei pruugi see tänapäevaste analüüsimeetodite, eriti tehisintellekti kasutamisel, olla piisav, et välistada kaudne isikutuvastus.

Pseudonümiseerimine on seevastu GDPR-is selgelt positsioneeritud riskide vähendamise vahendina, mitte mehhanismina, mis vabastaks töötaja andmekaitse nõuete täitmisest. GDPR artikkel 32 ning põhjendus 28 viitab, et pseudonümiseerimine aitab parandada töötlemise turvalisust ja vähendada andmesubjektide õiguste riive riski, kuid ei kõrvalda vajadust järgida teisi andmekaitsepõhimõtteid, sealhulgas minimaalsust, eesmärgipärasust ja läbipaistvust.⁶² Sellest tulenevalt võib pseudonümiseerimist käsitada kompromissina andmete kasutatavuse ja kaitse vahel, mis võimaldab andmete edasist töötlemist, kuid eeldab samaaegselt täiendavaid kontrolli- ja järelevalvemehhanisme.⁶³ Õiguslikus plaanis tähendab, et pseudonümiseerimine suurendab vastutava töötaja hoolsuskohustust, mitte ei vähenda seda.

⁶¹ Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 2010; vt ka A. Narayanan, V. Shmatikov, *Robust De-anonymization of Large Sparse Datasets*, 2008.

⁶² Euroopa Parlamendi ja nõukogu määrus (EL) 2016/679 2016. *Isikuandmete kaitse üldmäärus (General Data Protection Regulation)*. Euroopa Liidu Teataja L 119, 4.5.2016, lk 1–88.

⁶³ Regulation (EU) 2016/679, art 32 (töötlemise turvalisus) ja põhjendus 28 (pseudonümiseerimine kui riskide vähendamise meede), veebist kättesaadav: <https://gdpr-info.eu/art-32-gdpr/>; <https://gdpr-info.eu/recitals/no-28/>.

Anonümiseerimise ja pseudonümiseerimise eristus omandab tehisintellekti kontekstis erilise kaalu, kuna nende mõistete rakendamine mõjutab otseselt GDPR-i kohaldatavust ja seega ka arendajate ning kasutajate õiguslikku vastutust.⁶⁴ Samas on juhitud tähelepanu sellele, et praktikas võib nende kahe mõiste vaheline piir hägustuda, kuna tehnilised lahendused ei pruugi tagada täielikku pöördumatust ning andmete töötlemine toimub sageli keskkondades, kus andmete kombineerimine ja taaskasutus on vältimatu. See tekitab olukorra, kus kuigi andmete õiguslik kvalifikatsioon peab lõppastmes olema binaarne (kas isikuandmed või mitte ehk jah/ei loogika), võib anonümiseerimise tegelik vastavus sellele standardile olla tehniliselt ja kontekstuaalselt ebakindel ning sõltuda konkreetsetest kasutustingimustest ja mõistlikult eeldatavatest tuvastamisvõimalustest.

Anonümiseerimise ja pseudonümiseerimise peamine probleem seisnebki selles, et õigus eeldab nende puhul üsna selget ja püsivat eristust, kuid tegelikkuses ei ole see tehnoloogilises keskkonnas alati nii üheselt määratletav. Eriti LMM-ide puhul on süsteemid ajas kiirelt arenevad ning need õpivad andmetest viisil, mis võib hiljem võimaldada kaudset isiku tuvastamist või andmete osalist taastootmist. Lisaks ei sõltu anonümiseerimise tõhusus ainult kasutatavast tehnilisest lahendusest, vaid ka sellest, millises keskkonnas ja eesmärgil andmeid kasutatakse. Näiteks suletud ja rangelt kontrollitud süsteemides on risk üldjuhul väiksem, samas kui avatud ja dünaamilistes keskkondades, kus andmeid pidevalt täiendatakse ja kombineeritakse, võib tuvastamise risk oluliselt suurened.

Kuigi GDPR teeb selge õigusliku eristuse anonümiseerimise ja pseudonümiseerimise vahel, ei ole see eristus tehisintellekti kontekstis alati praktiliselt rakendatav. Pseudonümiseerimine jääb oma olemuselt riskide maandamise vahendiks, samas kui anonümiseerimine kui GDPR-i kohaldamisalast vabastav mehhanism on sageli raskesti saavutatav. See loob aluse järgmistes alapeatükkides käsitletavatele küsimustele, eeskätt seoses taasidentifitseerimise riskiga ja anonümiseerimise tegeliku tõhususega suurte keelemudelite puhul.

⁶⁴ Tupay, P. K., Ebers, M., Juksaar, J., & Kohv, K., *Is European data protection toxic for innovative AI? An Estonian perspective*, *Juridica International* 30 (2021), 193–206, DOI: 10.12697/JI.2021.30.12, veebist kättesaadav:

https://www.juridicainternational.eu/article_full.php?uri=2021_30_is_european_data_protection_toxic_for_innovative_ai_an_estonian_perspective.

3.2. „Pöördumatuse“ nõue ja taasidentifitseerimise risk AI kontekstis

Anonümiseerimise eelduseks GDPR-i tähenduses on nn pöördumatuse nõue, mille kohaselt peab isiku tuvastamine olema välistatud viisil, mis ei ole mõistlikult tõenäoline, arvestades kõiki asjakohaseid vahendeid ja asjaolusid. Kuigi GDPR ei sisalda anonümiseerimise ega pöördumatuse mõiste sõnaselget definitsiooni, tuleneb see nõue selgelt määruse põhjendusest 26, mille kohaselt tuleb tuvastatavuse hindamisel arvesse võtta kõiki objektiivseid tegureid, sealhulgas tehnoloogia arengut ja kolmandate isikute võimalusi. Sama lähenemist on süsteemselt arendanud ka Euroopa Liidu endine artikli 29 töörühm oma arvamuses anonümiseerimistehnike kohta, milles konsensuslikult leitakse, et anonümiseerimine eeldab isiku tuvastamise pöördumatut välistamist, mitte üksnes identifikaatorite eemaldamist.⁶⁵

Euroopa andmekaitsepraktikas on pöördumatuse mõiste tihedalt seotud identifitseeritavuse spektriga, mille kohaselt ei ole andmed kas „täielikult isikuandmed“ või „täielikult anonüümsed“, vaid paiknevad erinevatel vahepealsetel astmetel. Artikli 29 töörühm on leidnud, et anonümiseerimise tulemusel peab olema välistatud nii üksikisiku eristamine (*singling out*), andmekirjete omavaheline seostamine (*linkability*) kui ka uue isikut puudutava teabe järeldamine (*inference*).⁶⁶ Kui mõni neist riskidest püsib, ei saa anonümiseerimist pidada pöördumatuks isegi juhul, kui otsesed identifikaatorid on eemaldatud. Kuigi anonümiseerimine eeldab otseste identifikaatorite eemaldamist, ei ole see iseenesest piisav, et tagada pöördumatus, kuna identifitseeritavus võib säilida ka kaudsete tunnuste ja andmete kombineerimise kaudu. Sellest järeldub, et pöördumatus ei ole seotud konkreetse meetodi kasutamisega, vaid meetmete kumulatiivse mõjuga konkreetse andmestiku ja kasutuskonteksti puhul.

Tehisintellekti puhul ei piirdu andmete töötlemine üksnes nende otsese kasutamise või salvestamisega, vaid hõlmab ka mustrite, seoste ja latentsete representatsioonide õppimist. See tähendab, et isegi juhul, kui algandmed on formaalselt anonümiseeritud, võivad mudelid säilitada statistilist või semantilist teavet, mis võimaldab isikuid kaudselt tuvastada. Seetõttu on piisava lisainformatsiooni olemasolul võimalik anonümiseeritud andmekogumeid taasidentifitseerida, eriti kui neid kombineerida teiste andmeallikatega. Varasemad tööd

⁶⁵ Article 29 Data Protection Working Party, *Opinion 05/2014 on Anonymisation Techniques* (WP216), 10 April 2014, p-d 10–15, veebist kättesaadav: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

⁶⁶ *Ibid.*

deanonümiseerimise valdkonnas demonstreerisid nn mosaiikefekti, mille kohaselt üksikud, iseenesest mitteidentifitseerivad andmeelemendid võivad koosmõjus võimaldada konkreetse isiku tuvastamist.⁶⁷ Hilisemad uuringud on seda loogikat oluliselt edasi arendanud, näidates, et ka tänapäevaste anonümiseerimistehnike korral võib taasidentifitseerimise risk püsida kõrgena, eriti kui arvestada generatiivseid mudeleid ja kõrvalandmete kasvu.⁶⁸ See näitab selgelt, et anonümiseerimine ei ole püsiv ega absoluutne seisund, vaid ajas muutuv hinnang.

Suurte keelemudelite kontekstis lisandub sellele veel eripära, et mudelid võivad teatud tingimustel meelde jätta treeningandmete osi ning neid hiljem väljundina reprodutseerida. Uuringud on näidanud, et ebasoodsate tingimuste korral on võimalik mudelitest tuletada treeningandmeid, sealhulgas isikuandmeid, isegi kui see ei olnud arendaja algne eesmärk. Carlini jt uurimus nn treeningandmete eraldamise (training data extraction) kohta leiab, et mudelite keerukus ja mahukus suurendavad riski, et treeningandmete elemendid avalduvad mudeli väljundites.⁶⁹ Sellest tulenevalt ei ole võimalik täielikult välistada olukordi, kus isikuandmed muutuvad pärast anonümiseerimist uuesti tuvastatavaks. Praktikas muudab see pöördumatuse nõude täitmise suurte keelemudelite puhul äärmiselt keeruliseks, kui mitte sisuliselt saavutamatuks.

Euroopa andmekaitseasutused on asunud seisukohale, et anonümiseerimise hindamisel tuleb arvestada mitte ainult hetkeolukorda, vaid ka võimalikke tulevasi arenguid ning realistlikke taasidentifitseerimise võimalusi.⁷⁰ EDPB on leidnud, et anonümiseerimist ei saa pidada piisavaks, kui eksisteerib realistlik võimalus andmete taasidentifitseerimiseks, isegi juhul, kui see nõuab täiendavaid teadmisi, ressursse või tehnoloogilisi arenguid.⁷¹ Seisukoht viitab, et

⁶⁷ Narayanan, A., & Shmatikov, V. (2008). *Robust de-anonymization of large sparse datasets*. In *Proceedings of the 2008 IEEE Symposium on Security and Privacy* (pp. 111–125). IEEE. Veebist kättesaadav: <https://ieeexplore.ieee.org/document/4531148>

⁶⁸ Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). *Estimating the success of re-identifications in incomplete datasets using generative models*. *Nature Communications*, 10, 3069. Veebist kättesaadav: <https://www.nature.com/articles/s41467-019-10933-3>

⁶⁹ Carlini, N., Tramèr, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T., Song, D., Erlingsson, Ú., Oprea, A., & Raffel, C. (2021). *Extracting Training Data from Large Language Models*. In 30th USENIX Security Symposium (USENIX Security '21) (pp. 2633–2650). <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>

⁷⁰ Article 29 Data Protection Working Party, *Opinion 05/2014 on Anonymisation Techniques* (WP216), 10 April 2014, veebist kättesaadav: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

⁷¹ European Data Protection Board, *Report on the stakeholder event on anonymisation and pseudonymisation of 12 December 2025* (published February 2026), veebist kättesaadav: https://www.edpb.europa.eu/system/files/2026-02/edpb-report-stakeholder-event-anonymisation-pseudonymisation_en.pdf

anonümiseerimise standard on dünaamiline, mitte staatiline, ning sõltub otseselt tehnoloogilise arengu tasemest. Tehisintellekti kiire areng tähendab seega, et meetodid, mida täna peetakse piisavaks, võivad homme osutada ebapiisavaks.

Sarnane arusaam kajastub ka Eesti õiguskirjanduses. Näiteks märgivad Tupay jt, et tehisintellekti süsteemide arendamine seab GDPR-ist tulenevad andmekaitsemehhanismid, sealhulgas anonümiseerimise, märkimisväärse surve alla. AI-põhine andmetöötlus tugineb ulatuslikele ja kombineeritavatele andmetikele, mille puhul püsib kaudse tuvastatavuse risk ning õigusliku vastutuse piiritlemine muutub ebaselgeks.⁷² Autorite hinnangul omandab anonümiseerimise ja pseudonümiseerimise eristus AI-kontekstis erilise kaalu, kuna see mõjutab otseselt GDPR-i kohaldatavust ning arendajate ja kasutajate õiguslikku vastutust. Sellest tulenevalt ei saa anonümiseerimist käsitada pelgalt tehnilise protseduurina, vaid see eeldab laiaulatuslikku riskihinnangut, mis hõlmab nii andmete olemust, kasutuskonteksti kui ka süsteemi edasist arengut.

Tõden, et pöördumatuse nõude peamine probleem selles, et see eeldab anonümiseerimise lõplikkust ehk olukorda, kus isiku tuvastamine on igal juhul välistatud. Nagu eelpool leitud, ei ole see tehisintellekti kontekstis aga lõpuni kindel. Tegemist on süsteemidega, mille võimekus ajas areneb ning mille kasutusviise ei ole võimalik ette näha. Seetõttu ei ole anonümiseerimine kõigil juhtudel käsitatav absoluutselt pöördumatuna, vaid pigem kontekstipõhise seisundina, mille turvalisus sõltub ajas muutuvatest teguritest, näiteks tehnoloogia arengust, andmete kättesaadavusest ja nende kasutuskontekstist. Täna piisavalt anonüümseks peetud andmed ei pruugi seda olla tulevikus ning paraku ei sobitu asjade selline seis GDPR-i loogikaga, mis lähtub üsna jäigast eristusest anonüümsete andmete ja isikuandmete vahel.

Lisaks kerkib küsimus, kelle vaatenurgast pöördumatust hinnata. Kui hinnang tugineb üksnes vastutava töötleja käsutuses olevatele vahenditele, võib see alahinnata kolmandate isikute võimekust andmeid taasidentifitseerida. Kui aga arvesse võtta kõiki teoreetiliselt võimalikke arenguid, muutub anonümiseerimise absoluutne standard sisuliselt saavutamatuks. Seetõttu on

⁷² Tupay, P. K., Ebers, M., Juksaar, J., & Kohv, K., *Is European data protection toxic for innovative AI? An Estonian perspective*, *Juridica International* 30 (2021), 193–206, DOI: 10.12697/JI.2021.30.12, veebist kättesaadav:

https://www.juridicainternational.eu/article_full.php?uri=2021_30_is_european_data_protection_toxic_for_innovative_ai_an_estonian_perspective.

vajalik tasakaalustada realistlik ja hüpoteetiline riskihinnang, lähtudes mõistlikkuse ja proportsionaalsuse põhimõtetest, nagu on rõhutanud ka Euroopa Kohus hiljutises praktikas.⁷³

Võimaliku lahendusena võiks pöördumatuse nõuet tõlgendada riskipõhiselt, keskendudes mitte absoluutsele tuvastamatusele, vaid sellele, kas taasidentifitseerimise risk on konkreetsetes olukorras piisavalt madal ja kontrollitav. Nimetatud käsitlusviis haakub paremini GDPR-i üldise riskipõhise loogikaga ning võimaldab arvestada tehisintellekti ja andmetöötlemise dünaamilise iseloomuga. Samas ei tähenda see anonümiseerimise standardi sisulist leevendamist, vaid selle paindlikumat ja kontekstipõhist kohaldamist, lähtudes tegelikest riskidest ning kasutusolukorrast. Eelkõige tuleb anonümiseerimist käsitada mitte absoluutse garantiina, vaid riskijuhtimise vahendina, mille tõhusus sõltub konkreetsest kasutuskontekstist, tehnoloogilisest arengust ja rakendatud järelevalvemehhanismidest.

3.3. Pseudonümiseeritud andmete kasutamise õiguslikud piirid

Pseudonümiseerimine on GDPR-i raamistikus oluline tehniline ja organisatsiooniline meede, mille eesmärk on vähendada isikuandmete töötlemisega seotud riske, säilitades samal ajal andmete kasutatavuse. Erinevalt anonümiseerimisest ei too pseudonümiseerimine kaasa andmete väljajäämist GDPR-i kohaldamisalast, vaid kujutab endast riskimaandamise vahendit, mis mõjutab andmetöötlemise ohutaset ja sellest tulenevalt ka rakendatavate kohustuste sisu ja ulatust. Sellest tulenevalt on pseudonümiseeritud andmete töötlemine jätkuvalt allutatud kõigile andmekaitseõiguse aluspõhimõtetele, sealhulgas eesmärgipärasusele, minimaalsusele ja läbipaistvusele.⁷⁴

GDPR artikli 4 lõike 5 kohaselt tähendab pseudonümiseerimine isikuandmete töötlemist viisil, mille tulemusel ei ole andmeid võimalik seostada konkreetse isikuga ilma täiendava teabeta, tingimusel et see lisateave hoitakse eraldi ning on kaitstud asjakohaste tehniliste ja korralduslike meetmetega. Pseudonümiseerimine ei välista isiku tuvastatavust, vaid üksnes vähendab selle tõenäosust. Selle eesmärk on võimaldada isiku tuvastamist kontrollitud tingimustel (nt võtme või lisainfo abil), muutes samal ajal tuvastamise autoriseerimata isikutele praktiliselt võimatuks

⁷³ CJEU, *European Data Protection Supervisor v Single Resolution Board* (C-413/23 P), ECLI:EU:C:2025:645, p-d 92–104, veebist kättesaadav: <https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=CELEX%3A62023CJ0413>

⁷⁴ European Data Protection Board. (2025). *Guidelines 01/2025 on Pseudonymisation*. Veebist kättesaadav: https://www.edpb.europa.eu/system/files/2025-01/edpb_guidelines_202501_pseudonymisation_en.pdf

või ebamõistlikult keeruliseks. Näiteks võib haigla asendada patsientide nimed koodidega ning hoida nende koodide ja isikute vahelise seose eraldi turvalises andmebaasis. Sellisel juhul ei ole andmeid ilma täiendava teabeta võimalik konkreetse isikuga seostada, kuid vajaduse korral on isiku tuvastamine siiski võimalik, mistõttu on tegemist pseudonümiseerimisega, mitte anonümiseerimisega. Euroopa Kohtu praktika, sealhulgas kohtuasjas *Breyer v Germany* kujundatud lähenemine tuvastatavusele, kinnitab, et ka kaudne tuvastatavus on piisav, et käsitada andmeid isikuandmetena. Seetõttu ei saa pseudonümiseeritud andmeid käsitleda kui anonüümseid ning nende töötlemine peab vastama kõigile GDPR-i nõuetele.⁷⁵

Pseudonümiseeritud andmete kasutamise õiguslikke piire määratleb eeskätt eesmärgipärasuse põhimõte. Kuna pseudonümiseerimine ei muuda andmeid anonüümseks, ei anna see andmetöötlejale õigust kasutada andmeid uutel, ebamäärastel või algsest eesmärgist lahutatud eesmärkidel. Samuti ei kujuta pseudonümiseerimine endast õiguslikku vabastust, vaid on üks riskide maandamise meede, mis peab olema tihedalt seotud konkreetse ja õiguspärase töötlemise eesmärgiga.⁷⁶ Seega peab ka pseudonümiseeritud andmete kasutamine olema põhjendatud ning vastama proportsionaalsuse nõudele.

Lisaks seab pseudonümiseeritud andmete töötlemisele selged piirid minimaalsuse põhimõte. Nagu eelnevates peatükkides käsitletud, eeldab see põhimõte, et töödeldakse üksnes selliseid andmeid ja sellises mahus, mis on konkreetse eesmärgi saavutamiseks vältimatult vajalik. Pseudonümiseerimine ei vähenda iseenesest töödeldavate andmete hulka ega töötlemise ulatust, vaid üksnes alandab tuvastatavuse riski. Seetõttu ei saa pseudonümiseerimist käsitada minimaalsuse põhimõtte asendajana ega õigustusena ulatuslikumaks andmekogumiseks. Kirjanduses on seejuures osutatud, et pseudonümiseerimine võib eksliku tõlgenduse korral hoopis soodustada andmete liigset kogumist ja töötlemist, kui seda käsitatakse väärtelt piisava kaitsemeetmena.⁷⁷

LLM-ide arendamise kontekstis muutuvad pseudonümiseeritud andmete kasutamise piirid veelgi problemaatilisemaks. Sellistes süsteemides ei ole andmetöötlus lineaarne ega täielikult

⁷⁵ CJEU, *Patrick Breyer v Bundesrepublik Deutschland* (C-582/14), ECLI:EU:C:2016:779, p-d 41–49, veebist kättesaadav: <https://eur-lex.europa.eu/legal-content/EN/SUM/?uri=CELEX:62014CJ0582>.

⁷⁶ European Data Protection Board, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default* (2020)

⁷⁷ Purtova, N., *The law of everything: Broad concept of personal data and future of EU data protection law*, *Law, Innovation and Technology* 10(1), 2018, lk 40–81, <https://doi.org/10.1080/17579961.2018.1452176>

ette prognoositav ning pseudonümiseeritud andmed võivad olla sisendiks mudelitele, mis suudavad tuletada uusi seoseid või järeldusi, suurendades seeläbi kaudse tuvastatavuse riski. Lisaks võib tekkida olukordi, kus mudelite väljundid sisaldavad isikuandmeid või nendega võrreldavat infot, hoolimata sellest, et sisendandmed olid pseudonümiseeritud.⁷⁸ Sellest tulenevalt tuleb järeldada, et pseudonümiseerimine ei pruugi tehisintellekti kontekstis tagada andmesubjektide õiguste piisavat ja tõhusat kaitset, kuna see ei kõrvalda isikuandmete kaudse tuvastatavuse ega andmete väljundipõhise taastootmise riski.

Tehisintellekti arendamise kontekstis omandab pseudonümiseerimine eeskätt riskimaandamise ja vastutuse ümberjaotamise funktsiooni, mitte aga regulatiivse vabastuse tähendust. Pseudonümiseerimise roll ei seisne mitte andmete anonüümseks muutmises, vaid selles, et see aitab vähendada andmetöötlusega seotud riske ja toetab GDPR-i põhimõtete järgimist. Samas, pseudonümiseeritud andmete kasutamine eeldab selgelt määratletud töötlemise eesmärgi ning täiendavate tehniliste ja organisatsiooniliste kaitsemeetmete rakendamist, kuna vastasel juhul jääb alles oht, et andmete kombineerimise või edasise töötlemise kaudu võib taastuda isikuandmete tuvastatavus.⁷⁹ Nõustun seisukohaga, et pseudonümiseerimine ei tähenda GDPR-i nõuete kõrvale jätmist, vaid on osa laiemast riskipõhisest lähenemisest andmekaitsele, mis eeldab vastutuse ja andmetöötlusriskide järjepidevat läbimõtlemist kogu tehisintellekti süsteemi elukaare vältel. Eriti tehisintellekti kontekstis võib pseudonümiseerimine luua näilise turvatunde, kuigi andmete tegelik kasutusviis ja nendest tuletatavad järeldused ei ole alati täielikult kontrollitavad. Kirjeldatud olukord võib omakorda viia andmetöötlejate kohustuste alahindamiseni ning raskendada tõhusa järelevalve teostamist.

Pseudonümiseerimise kasutamist tuleks aga käsitada rangemalt ning siduda see selgete reeglitega, nagu kohustuslikud riskihinnangud, kindlad töömeetodid ja perioodilised ülevaatused. Oluline on, et need ei keskenduks üksnes tehnilise lahenduse olemasolule, vaid hindaksid ka seda, milline on pseudonümiseerimise tegelik mõju tuvastatavuse riskile konkreetses kasutuskontekstis.

⁷⁸ Carlini, N., Tramèr, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T., Song, D., Erlingsson, Ú., Oprea, A., & Raffel, C. (2021). *Extracting Training Data from Large Language Models*. In *30th USENIX Security Symposium (USENIX Security 2021)*, pp. 2633–2650. Veebist kättesaadav: <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>

⁷⁹ Tupay, P. K., Ebers, M., Juksaar, J., & Kohv, K., Is European data protection toxic for innovative AI? An Estonian perspective, *Juridica International* 30 (2021), 193–206, DOI: 10.12697/JI.2021.30.12, veebist kättesaadav: https://www.juridicainternational.eu/article_full.php?uri=2021_30_is_european_data_protection_toxic_for_innovative_ai_an_estonian_perspective.

3.4. Anonümiseerimise roll andmete minimaalsuse alternatiivina

Anonümiseerimist on andmekaitseõiguses sageli käsitatud võimaliku lahendusena olukordades, kus andmete minimaalsuse põhimõtte range järgimine osutub praktiliselt keeruliseks. Kuna anonümiseeritud andmed ei kuulu üldjuhul GDPR-i kohaldamisalasse, võib esmapilgul jääda mulje, et anonümiseerimine võimaldab ületada minimaalsuse piiranguid ning töödelda mahukaid andmekogumeid ilma andmekaitseõiguslike takistusteta. Selline käsitus on eriti levinud tehisintellekti, sealhulgas LLM-ide kontekstis, kus süsteemide tõhus toimimine eeldab ulatuslikku ja mitmekesist andmestikku.⁸⁰ Siiski ei saa anonümiseerimist pidada andmete minimaalsuse põhimõtte asenduseks, vaid üksnes piiratud ulatusega täiendavaks meetmeks.

Esiteks, andmete minimaalsuse ja anonümiseerimise normatiivsed eesmärgid on olemuslikult erinevad. Minimaalsuse põhimõtte keskendub küsimusele, millises ulatuses on andmete kogumine ja töötlemine konkreetse eesmärgi saavutamiseks vajalik ja proportsionaalne. Anonümiseerimine seevastu tegeleb andmete tuvastatavuse kõrvaldamisega pärast seda, kui andmed on juba kogutud. Seega, ei lahenda anonümiseerimine küsimust andmete kogumise vajalikkusest, vaid püüab vähendada andmete edasisest kasutamisest tulenevaid riske. Samuti ei saa anonümiseerimist käsitada universaalse lahendusena, kuna selle tõhusus sõltub alati konkreetsetest tehnilistest, organisatsioonilistest ja kontekstuaalsetest asjaoludest ning anonümiseerimine ei toimi iseenesest väljaspool laiemat andmekaitseraamistikku.⁸¹

Kirjanduses on leitud, et anonümiseerimine võib praktikas viia nn funktsionaalse minimaalsuse nõrgenemiseni. Kui andmeid käsitatakse anonüümsetena, võib see soodustada nende laiemat kogumist ja edasist kasutamist, kuna otsene andmekaitseõiguslik piirang näib puuduvat. Andmekaitseõiguse laienev kohaldamisala ja samal ajal anonümiseerimisele tuginemine võivad paradoksaalselt viia olukorrani, kus andmete kogumise maht suureneb, ilma et andmesubjektide tegelik risk oluliselt väheneks.⁸² Need riskid avalduvad eeskätt andmete kogumise etapis, kuid ulatuvad ka nende edasisse kasutamisse ja kombineerimisse, suurendades seeläbi kaudse

⁸⁰ Article 29 Data Protection Working Party. (2014). *Opinion 05/2014 on Anonymisation Techniques* (WP216). Adopted on 10 April 2014. Veebist kättesaadav: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

⁸¹ *Ibid.*

⁸² Purtova, N., *The law of everything: Broad concept of personal data and future of EU data protection law*, *Law, Innovation and Technology* 10(1), 2018, lk 40–81, <https://doi.org/10.1080/17579961.2018.1452176>

tuvastatavuse võimalust. Seega ka anonümiseerimine ei toimi andmete minimaalsuse põhimõtte asendajana, vaid võib teatud tingimustel seda hoopis sisuliselt õõnestada.

Tehisintellekti kontekstis muutub see probleem eriti teravaks. LLM-ide arendamisel kasutatakse sageli väga suuri andmekogumeid, mille puhul eeldatakse, et anonümiseerimine võimaldab vältida andmekaitseõiguslikke piiranguid. Samas on uuringud näidanud, et isegi anonümiseeritud andmete puhul võib esineda taasidentifitseerimise oht, näiteks andmete kombineerimise, mustrite tuvastamise või mudelite väljundite kaudu. On leitud, et piisava lisainformatsiooni olemasolul võivad ka anonümiseeritud andmekogumid osutada haavatavaks kaudse tuvastatavuse suhtes.⁸³ Tuginedes eeltoodule, saab seada kahtluse alla anonümiseerimise kui iseseisva ja piisava kaitsemehhanismi.

Kuigi Eestis ei ole seni avalikult dokumenteeritud konkreetseid juhtumeid, kus anonümiseeritud andmetel treenitud tehisintellekti süsteem oleks toonud kaasa isikuandmete ilmnemise mudeli väljundites, teadvustatakse seda riski järelevalvepraktikas. Eesti andmekaitseasutuse ja keeletehnoloogia valdkonna käsitlused kinnitavad, et anonümiseerimine ei pruugi tehisintellekti kontekstis olla püsiv ega piisav kaitsemehhanism, kuna mudelite võime tuletada uusi seoseid võib kaasa tuua isikute kaudse tuvastamise ka formaalselt anonümiseeritud andmete puhul.⁸⁴

Euroopa Liidu andmekaitseraamistikust ja järelevalvepraktikast tuleneb, et anonümiseerimine ei vabasta andmetöötajat andmetöötamise vajalikkuse ja proportsionaalsuse hindamisest. GDPR käsitleb anonümiseerimist ühe võimaliku riskimaandamise meetmena, mitte andmekaitsepõhimõtete asendusena. Sama lähenemist peegeldab ka Eesti andmekaitse järelevalvepraktika, mille kohaselt tuleb anonümiseerimise piisavust hinnata sisuliselt ning arvestada ka kaudse tuvastatavuse ja tehnoloogilise arengu riskiga.⁸⁵

⁸³ Narayanan, A., & Shmatikov, V., *Robust de-anonymization of large sparse datasets*, *IEEE Symposium on Security and Privacy*, 2008, lk 111–125, <https://doi.org/10.1109/SP.2008.33>

⁸⁴ Andmekaitse Inspeksioon. (2024). *Tähelepanekuid andmekaitse ja tehisintellekti (AI) seostest ja arengutest*. In: *Andmekaitse Inspeksiooni aastaraamat 2024*. Veebist kättesaadav: <https://aastaraamat.aki.ee/aastaraamat-2024/tahelepanekuid-andmekaitse-ja-tehisintellekti-ai-seostest-ja-arengutest>

⁸⁵ Andmekaitse Inspeksioon. (2021–2022). *Lühiseletused ja seisukohad anonümiseerimise ning kaudse tuvastatavuse kohta*. In: *Andmekaitse Inspeksiooni aastaraamat 2021–2022*. Veebist kättesaadav: <https://aastaraamat.aki.ee/aastaraamat-2021-2022/luhiseletused>

Tõden, et anonümiseerimise kui minimaalsuse asenduse suurim probleem on see, et see nihutab tähelepanu valesse kohta. Andmete minimaalsuse põhimõtte sunnib juba ette küsima, kas ja kui palju andmeid on üldse vaja koguda. Anonümiseerimine seevastu eeldab, et andmed on juba kogutud, ning keskendub sellele, kuidas hiljem võimalikke riske vähendada. See ei ole kooskõlas GDPR-i loogikaga, mille eesmärk on ennetada liigset andmetöötlust juba planeerimise ja algfaasi tasandil. Samas ei saa anonümiseerimise tähtsust täielikult eitada. Teatud olukordades võib see olla vajalik ja kasulik vahend, näiteks teadusuuringute või andmepõhise innovatsiooni toetamisel, kus ilma anonümiseerimiseta ei oleks andmete kasutamine võimalik. Ka neil juhtudel tuleb aga arvestada, et anonümiseerimine ei kõrvalda kõiki riske ning selle sobivus sõltub alati konkreetsest tehnilisest ja õiguslikust kontekstist.

Seetõttu, ei vabasta anonümiseerimine andmetöötlejat kohustusest hinnata andmete kogumise vajalikkust ega taga automaatselt andmesubjektide õiguste kaitset. Tehisintellekti kontekstis, kus andmete kogumine on süsteemide toimimise põhiline eeldus, on eriti oluline vältida olukorda, kus anonümiseerimine muutub vahendiks, millega õigustatakse andmete minimaalsuse põhimõtte sisulist eiramist.

3.5. Õiguslikud ja praktilised probleemid LLM-ide treeningandmetes

LLM-ide arendamine tugineb ulatuslikele treeningandmetikele, mille kogumine, töötlemine ja hilisem kasutamine toovad kaasa mitmetasandilisi õiguslikke ja praktilisi probleeme. Nagu eelnevates alapeatükkides on käsitletud, ei ole traditsioonilised andmekaitsemehhanismid, eeskätt anonümiseerimine ja pseudonümiseerimine, LLM-ide kontekstis piisavad, et tagada GDPR-i põhimõtete tegelikku järgimist. Probleemid koonduvad eelkõige treeningandmete päritolu ja õiguspärasuse, töötlemise läbipaistvuse, andmete minimaalsuse ning püsiva tuvastatavuse riski ümber.

Esiteks tõusetub keskse õigusliku küsimusena treeningandmete päritolu aspekt. LLM-ide arendamisel kasutatakse sageli automatiseeritud meetoditega kogutud veebipõhiseid andmekogumeid, mille sisu ja ulatus ei ole arendajatele täielikult teada. Selliste andmetike puhul puudub sageli selge ülevaade, kas ja millisel määral need sisaldavad isikuandmeid, millistest allikatest andmed pärinevad ning millistel tingimustel need algselt avalikustati. GDPR eeldab aga, et vastutav töötleja suudab tuvastada töödeldavate andmete liigi ja päritolu ning hinnata nende töötlemise õiguspärasust, sealhulgas määratleda sobiva õigusliku aluse. Kui

selline kontroll ja teadlikkus puudub, muutub ka õiguspärase töötlemise aluse ehk õigustatud huvi kasutamine sisuliselt problemaatiliseks.⁸⁶

Teiseks seab LLM-ide treeningandmete kasutamine tõsise väljakutse läbipaistvuse põhimõtte rakendamisele. GDPR näeb ette, et andmesubjektidele tuleb anda selget ja arusaadavat teavet nende isikuandmete töötlemise kohta. LLM-ide puhul on aga praktiliselt võimatu teavitada kõiki isikuid, kelle andmeid võib olla treeningandmetesse kaasatud, eriti juhul, kui andmed pärinevad avalikest, kuid struktureerimata veebiallikatest.⁸⁷ Lisaks ei ole sageli võimalik tuvastada konkreetseid andmesubjekte, mis muudab keeruliseks ka nende õiguste, näiteks juurdepääsu-, parandamis- või kustutamisoiguse, tegeliku teostamise. Mistõttu läbipaistvuse põhimõtte ei ole LLM-ide kontekstis mitte üksnes raskesti rakendatav, vaid võib oma algses tähenduses osutuda sisuliselt ebapiisavaks.⁸⁸

Kolmandaks tekib oluline ebakõla andmete minimaalsuse põhimõttega. LLM-ide arendamise loogika lähtub eeldusest, et suuremad ja mitmekesisemad andmekogumid parandavad mudelite kvaliteeti ja üldistamisvõimet. See lähenemine on aga otseses vastuolus minimaalsuse põhimõttega, mis nõuab andmete kogumise ja töötlemise piiramist üksnes vajalikuni. Praktikas on aga keeruline määratleda, milline andmehulk on LLM-ide puhul vajalik, kuna mudelite toimivus ei sõltu üksnes andmete mahust, vaid ka nende kvaliteedist, esitusviisist ja töötlemismeetoditest. Selline ebamäärasus muudab minimaalsuse põhimõtte sisulise järgimise LLM-ide kontekstis eriti problemaatiliseks.⁸⁹

Neljas probleem puudutab anonümiseerimise ja pseudonümiseerimise tegelikku tõhusust. Kuigi neid meetmeid kasutatakse sageli riskide vähendamiseks, on teaduskirjanduses leitud, et LLM-id võivad teatud tingimustel meelde jätta treeningandmete osi ning neid hiljem väljundites taasesitada. Carlini jt on empiiriliselt tõendanud, et suured keelemudelid võivad päringute kaudu reprodutseerida treeningandmetes sisalduvat teavet, sealhulgas

⁸⁶ Euroopa Parlamendi ja nõukogu määrus (EL) 2016/679 (GDPR), art 5 (põhimõtted) ja art 6 (töötlemise seaduslikkus), veebist kättesaadav: <https://gdpr-info.eu/art-5-gdpr/>; <https://gdpr-info.eu/art-6-gdpr/>.

⁸⁷ European Data Protection Board (EDPB), pressiteade ja viide arvamusele AI-mudelite kohta, 18.12.2024, veebist kättesaadav: https://www.edpb.europa.eu/news/news/2024/edpb-opinion-ai-models-gdpr-principles-support-responsible-ai_et.

⁸⁸ GDPR art 12 (teavitamise ja suhtluse läbipaistvus), art 13 (andmed saadakse andmesubjektilt) ja art 14 (andmed saadakse muust allikast), veebist kättesaadav: <https://gdpr-info.eu/art-12-gdpr/>; <https://gdpr-info.eu/art-13-gdpr/>; <https://gdpr-info.eu/art-14-gdpr/>.

⁸⁹ Purtova, N., *The law of everything: Broad concept of personal data and future of EU data protection law*, *Law, Innovation and Technology* 10(1), 2018, lk 40–81, DOI: <https://doi.org/10.1080/17579961.2018.1452176>

isikuandmeid.⁹⁰ See tähendab, et isegi eelnevalt töödeldud andmete puhul ei ole võimalik täielikult välistada isikuandmete hilisemat avaldumist, mis seab kahtluse alla anonümiseerimise kui lõpliku ja pöördumatu kaitsemeetme.

Lisaks tuleb arvestada, et LLM-ide andmetöötlus ei piirdu üksnes treeningfaasiga. Mudelid on loodud pidevaks kasutamiseks ja edasiarendamiseks ning võivad genereerida väljundeid, mis sisaldavad isikuandmeid või võimaldavad isikute kaudset tuvastamist. Seega ei avaldu andmekaitseõiguslikud riskid mitte ainult andmete kogumisel, vaid kogu süsteemi elutsükli vältel, hõlmates nii arendamist, kasutuselevõttu kui ka hilisemat kasutamist. Seda elutsükli põhise riski on rõhutanud ka Euroopa Andmekaitseinspektsiooni, mille kohaselt tuleb tehisintellekti mudelite puhul hinnata anonüümsust ja isikuandmete töötlemist ka väljundite tasandil.⁹¹

Kuigi Eesti kohtupraktikas ei ole seni lahendatud vaidlusi, mis puudutaksid otseselt LLM-ide treenimist või generatiivse AI treeningandmete kogumist, on algoritmilist skoorimist, profiilianalüüsi ning avalikest allikatest pärineva teabe uuskasutust käsitlevad lahendid sisuliselt heaks lähtekohaks. Need aitavad siduda käesolevas alapeatükis käsitletud treeningandmete probleemid (päritolu, läbipaistvus, minimaalsus ja elutsükli põhised riskid) konkreetsete õiguslike kriteeriumidega: (i) millal tekivad töötamise käigus uued, tuletatud isikuandmed (inferentsid), (ii) kuidas hinnata avaliku teabe edasist kasutust uuel eesmärgil ning (iii) miks tuleb proportsionaalsust hinnata ka andmetöötluse kogumõju põhjal ja mitte ainult üksikandme tasandil.

Eeltoodut täpsustab Euroopa Kohtu praktika, millele viitab ka Andmekaitse Inspektsiooni (AKI) aastaraamatu 2025 ülevaade „Olulised kohtuasjad“.⁹² Esiteks on LLM-ide nn musta kasti (black box) ja ärisaladuse argumendi puhul oluline lahend C-203/22 (*CK v Dun & Bradstreet Austria*), milles kohus leidis, et andmesubjektil peab olema õigus saada sisukat teavet automatiseeritud töötamise loogika kohta ning et pelgalt üldsõnaline või liiga abstraktne kirjeldus

⁹⁰ Carlini, N. jt, *Extracting Training Data from Large Language Models*, 30th USENIX Security Symposium (USENIX Security '21), 2021, veebist kättesaadav: <https://www.usenix.org/system/files/sec21-carlini-extracting.pdf>

⁹¹ European Data Protection Board (EDPB), pressiteade ja viide arvamusele AI-mudelite kohta (anonüümsuse hindamine juhtumipõhiselt ning mudeli kaudu päringutega isikuandmete saamise risk), 18.12.2024, veebist kättesaadav:

https://www.edpb.europa.eu/news/news/2024/edpb-opinion-ai-models-gdpr-principles-support-responsible-ai_et

⁹² Andmekaitse Inspektsioon, *Aastaraamat 2025*, „Olulised kohtuasjad“, kättesaadav: <https://www.aki.ee>

ei pruugi täita läbipaistvuse eesmärki. Kohus rõhutas seejuures, et kuigi vastutav töötleja võib tugineda ärisaladuse kaitsele, ei saa seda kasutada andmesubjekti õiguste sisuliseks piiramiseks. See tähendab, et ärisaladusele viitamine ei vabasta kohustusest selgitada otsuste tegemise aluseks olevaid põhimõtteid, loogikat ja peamisi tegureid arusaadaval viisil. Vajadusel tuleb leida tasakaal, näiteks avaldades piisava selgituse ilma tundlikke tehnilisi detaile täielikult paljastamata.⁹³ Selline lähenemine on eriti oluline tehisintellekti süsteemide puhul, kus läbipaistvuse nõue ei tähenda mitte algoritmi täielikku avalikustamist, vaid seda, et andmesubjekt saaks aru, kuidas ja milliste tegurite alusel tema kohta järeldusi tehakse.

Pseudonümiseerimise ja tuvastatavuse hindamisel oluline lahend C-413/23 P (*EDPS v Single Resolution Board*), kus on asutud seisukohale, et tuvastatavus on kontekstipõhine ning sõltub muu hulgas sellest, millised mõistlikult tõenäolised vahendid on konkreetse saaja või osapoole käsutuses.⁹⁴ LLM-ide treeningandmete kontekstis toetavad need lahendid järeldust, et nii läbipaistvust kui ka (pseudo)anonüümsust tuleb hinnata ökosüsteemiüleselt, arvestades iga asjaosalise, eelkõige vastutava töötleja, volitatud töötleja ning võimalike kolmandate isikute, tegelikke rolle, juurdepääse ja realistlikke riskistsenaariume.

Inferentsandmete ehk järeldatud andmete ja uue andmekihi tekkimist illustreerib Tallinna Haldus- ja Ringkonnakohtu praktikas käsitletud krediidi- ja maineskooride vaidlus (haldusasi nr 3-21-1749-veebilehed Kredix, Register ja Storybook). Kohus tuvastas, et olemasolevatest andmetest automaatselt loodud skoorid ja hinnangud ei ole pelgalt algandmete tehniline kuvamine, vaid isiku kohta käiv uus järelduslik teave, mis võib mõjutada tema mainet ja majanduslikku positsiooni.⁹⁵ LLM-ide treeningandmete kontekstis aitab see mõista, miks mudeli õpitud esitused ja selle väljundis esitatavad väited inimese kohta ei ole õiguslikult neutraalsed. Kui väljundi põhjal on võimalik inimest tuvastada või teda hinnata, võib tegemist olla isikuandmetega. Seetõttu ei tohiks treeningandmete õiguspärasuse ja minimaalsuse arutelu piirduda üksnes sisendandmete nimekirja või anonümiseerimisega, vaid peab hõlmama ka prognoositavat väljundiriski (sh võimalust, et mudel taastoodab või konstrueerib hinnangulisi järeldusi).

⁹³ C-203/22, *CK v Dun & Bradstreet Austria GmbH*, EU:C:2023:920, 7.12.2023.

⁹⁴ C-413/23 P, *EDPS v Single Resolution Board*, Euroopa Kohtu otsus 4.09.2025, EU:C:2025:645

⁹⁵ Tallinna Ringkonnakohus, 28.05.2024, haldusasi nr 3-21-1749, p-d 11–12 (algoritmiliselt arvutatud skoor kui isikuandmed).

Avalikest allikatest pärinevate andmete uuskasutuse piire käsitles Tallinna Halduskohtu 9. jaanuari 2025. aasta otsus eraisiku ja Andmekaitse Inspektsiooni vaidluses (sotsiaalmeedias avaldatud video ja heli töötlemine ning levitamine). Kohtu põhisoõnum seisnes selles, et avalikus ruumis avaldatud teave ei muutu seetõttu automaatselt piiranguteta kasutatavaks.⁹⁶ Otsustavaks jääb uue töötlemise eesmärk ja selle seos algse avaldamise kontekstiga. LLM-ide treeningpraktikate puhul, kus andmeid kogutakse massiliselt veebist (sh sotsiaalmeediast ja foorumitest), aitab see lahend selgitada väite piire, mille kohaselt võib avalikult kättesaadavaid veebiandmeid vabalt kasutada. Pelgalt andmete kättesaadavus ei tähenda, et neid võib kasutada mudeli üldotstarbeliseks treenimiseks, ilma et oleks sisuliselt hinnatud läbipaistvuse ja eesmärgipärasuse nõuete täitmist. Küsimus on selles, kas andmesubjekt saab mõistlikult ette näha, et tema postitust kasutatakse treeningandmetena, ning kas tal on tegelik kontroll, kui andmeid kasutatakse algsest kontekstist erineval eesmärgil.

Täiendavalt on minimaalsuse ja proportsionaalsuse mõju hindamisega sisuliselt lähedane Tallinna Halduskohtu praktika maksehäirete avaldamise asjas (haldusasi nr 3-23-2208).⁹⁷ Kohus analüüsis, kas standardiseeritud ja süsteemne andmete avaldamine on proportsionaalne ning kas andmetöötluse kogumõju ei põhjusta isikule ülemäärast kahju. LLM-ide treeningandmete puhul on see võrdlus oluline, sest risk ei pruugi tekkida ühest konkreetsest andmest, vaid sellest, kui palju andmeid kogutakse ja kui laialdaselt neid hiljem kasutatakse. Isegi kui ühe inimese üksikandme kasutamine tundub väike asi, võib suurte andmehulkade, mudeli laialdase kasutuse ja hilisema ümberkasutamise tõttu tekkida märkimisväärne mõju, näiteks andmeleke, mainekahju või diskrimineerivad järeldused. Seega, minimaalsuse hindamine LLM-ide puhul peab nihkuma üksikute andmeelementide kontrollilt süsteemsele tasandile, hõlmates realistlikke riskistsenaariume ja kogu elutsükli.

Euroopa Kohtu praktika asjas *CK v Dun & Bradstreet Austria* lisab LLM-ide konteksti jaoks olulise rõhuasetuse: kui töötlemine põhineb automatiseeritud otsustusprotsessil ja selle tulemused mõjutavad isikut, ei piisa üksnes väitest, et mudeli või andmestiku detailid on ärisaladus või tehniliselt keerukad. Läbipaistvus peab olema sisuline sellisel määral, et andmesubjekt (või tema esindaja) saaks reaalselt hinnata, milliseid andmeid kasutati, milliseid põhimõtteid rakendati ja milline oli järelduse kujunemise üldine menetlus.⁹⁸ See on oluline ka

⁹⁶ Tallinna Halduskohus, 9.01.2025, haldusasi nr 3-24-2146, p 18 jj.

⁹⁷ Tallinna Halduskohus, 30.01.2024, haldusasi nr 3-23-2208, p 14.

⁹⁸ C-203/22, *CK v Dun & Bradstreet Austria GmbH*, EU:C:2023:920, 7.12.2023.

LLM-ide treeningandmete vaates, kui treeningandmete päritolu ja filtreerimise loogika jäävad täielikult varjatuks, suureneb risk, et läbipaistvus taandub deklaratsiooniks ning andmesubjekti õigused muutuvad praktikas teostamatuks.

Lisaks haakub kogumõju käsitlus LLM-ide treeningandmete puhul ka identifitseeritavuse küsimusega: EK asjas C-413/23 P (*EDPS v SRB*) valguses ei saa (pseudo)anonüümsust hinnata ainult ühe osapoole (nt mudeli arendaja) vaatenurgast, vaid arvestada tuleb, millised tuvastamisvahendid on mõistlikult tõenäoliselt olemas ka teistel ahela osalistel (nt teenusepakkuja, tehniline vahendaja, klient, järelevalve).⁹⁹ Praktikas tähendab see, et isegi kui arendaja käsitleb treeningandmeid kui pseudonümiseeritud või anonüümse riskiga, võib tuvastatavus taastuda andmevoogude ristumisel või väljundite kaudu teise osapoole käes. Seetõttu tugevdab Euroopa Kohtu lähenemine järeldust, et minimaalsuse ja turvameetmete sisuline hindamine peab hõlmama kogu ökosüsteemi ning realistlikku vaadet sellele, kuidas andmeid (või mudeli väljundeid) tegelikult kasutatakse.

Eesti kohtupraktika toetab käesolevas alapeatükis tehtud järeldusi kahel tasandil. Esiteks kinnitab see, et automaatselt loodud järeldused ja hinnangud võivad ise olla isikuandmed ning sellest tulenevalt ei saa LLM-ide puhul piirduda üksnes toorandmete päritolu ja anonümiseerimise kontrolliga, vaid hinnata tuleb ka seda, milliseid isikut puudutavaid väiteid mudel võib realistlikult genereerida. Teiseks rõhutab kohtupraktika, et avalikest allikatest kogutud andmete uuskasutus eeldab jätkuvalt eesmärgipärasuse ja läbipaistvuse nõuete sisulist hindamist. Samuti andmete töötlemise vajalikkuse ja ulatuse põhjendatust, eriti olukorras, kus andmetöötlus on massiline ning selle mõju võib ilmned mudeli treenimisel, kasutamisel ja väljundites.

Järelikult seisneb LLM-ide treeningandmete puhul üldine probleem selles, et andmete päritolu, koosseis ja kasutamise alus ei ole piisavalt läbipaistvad ega hinnatavad andmete minimaalsuse põhimõtte seisukohalt. Andmeid kogutakse massiliselt ja eri allikatest, nende päritolu ning kasutusõigus ei ole alati selgelt tuvastatav ning andmesubjekti jaoks on selline uuskasutus enamasti ettenägematu. Sellises olukorras ei toimi anonümiseerimine või pseudonümiseerimine iseseisva lahendusena, vaid kujutab endast üht võimalikku riskimaandamise meetet, millele peab lisanduma andmete töötlemise õigusliku aluse, vajalikkuse ja ulatuse sisuline

⁹⁹ C-413/23 P, *EDPS v Single Resolution Board*, Euroopa Kohtu otsus 4.09.2025, EU:C:2025:645

põhjendamine. Seetõttu eeldab GDPR-iga kooskõla LLM-ide treeningandmete puhul eelkõige piisavalt läbipaistvat ja auditeeritavat dokumentatsiooni, mis kirjeldab muu hulgas kasutatavaid andmeallikaid, andmekategooriaid ja andmete töötlemisel rakendatavaid filtreid, ning realistlikku riskihindamist, mis arvestab ka mudeli väljundite ja hilisema kasutusega seotud mõjusid.

KOKKUVÕTE

Käesoleva magistritöö eesmärk oli selgitada, millisel määral ja millistel tingimustel on võimalik ühildada GDPR-ist tulenevad andmekaitsepõhimõtted, eeskätt andmete minimaalsuse põhimõtte ning anonümiseerimise ja pseudonümiseerimisega seotud nõuded, tehisintellekti, eriti LLM-ide arendamise ja kasutamise praktikatega. Uurimisprobleem väljendub normatiivse dilemmana ehk kuidas rakendada andmekaitseõiguse ennetavat ja piiravat loogikat olukorras, kus andmetöötlus on mahukas, dünaamiline ja piiratud läbipaistvusega ning kus andmete väärtus ja kasutusviisid selguvad sageli alles hilisemas etapis.

Töö läbiv seisukoht on, et GDPR-i põhimõtted ei kaota tehisintellekti kontekstis oma kehtivust ega muutu pelgalt formaalseteks nõueteks, millel puuduks praktiline tähendus. Samas näitab analüüs, et nende põhimõtete rakendamine LLM-ide arendamisel ja kasutamisel on oluliselt keerukam. Eelkõige tuleneb see asjaolust, et andmeid kogutakse massiliselt ja mitmekesistest allikatest, nende päritolu ja kasutus ei ole alati täielikult läbipaistvad ning andmete edasine kasutus ei pruugi olla algselt selgelt piiritletav. Sellises olukorras ei ole andmetöötlus käsitatav üksnes eelnevalt määratletud eesmärgi täitmisena, vaid hõlmab ka andmestikust tuletatavaid üldistusi, mille mõjud võivad avalduda hiljem mudeli väljundites ja edasises kasutuses. Seega ei ole küsimus ainult praktilistes raskustes, vaid ka sügavamas vastuolus andmekaitseõiguse ja tehisintellekti arenduse vahel.

LLM-ide kontekstis ei ole andmete vajalikkust enam võimalik sama selgelt ja täpselt mõõta kui varasemates käsitlustes. Traditsiooniline minimaalsuse käsitlus eeldab, et vajalik andmehulk on määratletav enne töötlemist, kuid LLM-ide puhul selgub andmete tegelik lisaväärtus sageli alles töötlemise käigus või tagantjäreli. Mudeli toimivus ei sõltu üksnes andmemahust, vaid ka andmekvaliteedist, andmete päritolust, jaotusest, duplikaatidest, filtreerimisest ja kasutusjuhtudest. Seetõttu ei saa andmete minimaalsuse põhimõtte tegelikku järgimist usaldusväärselt tagada üksnes andmete hulga piiramisega. Pigem tuleb minimaalsust LLM-ide puhul rakendada koos vajalikkuse, proportsionaalsuse ja riskihindamisega, hinnates, kas kasutatavad andmed on eesmärgi suhtes põhjendatud ning millised riskid nende töötlemisega kaasnevad ja kuidas neid maandatakse. Selline käsitlus ei tühistata minimaalsust kui normatiivset kohustust, vaid kohandab selle rakendamist LLM-ide andmemahuka ja dünaamilise arendusloogikaga. Oluline ei ole enam üksnes see, kui palju andmeid töödeldakse, vaid millise mõjuga on andmetöötlus ning kas sellega kaasnevad riskid on põhjendatavad ja kontrollitavad

Läbipaistvuse osas selgus, et tavapärane lähenemine, mis eeldab iga andmesubjekti eraldi teavitamist, ei sobitu LLM-ide treeningandmete konteksti. Kui andmeid kogutakse väga paljudest ja mitmekesistest allikatest, ei ole realistlik kõiki puudutatud isikuid tuvastada ega neid klassikalises tähenduses teavitada. Lisaks on keeruline selgitada mudeli toimimist, kuna LLM-id ei tegutse ette määratud sisendi-väljundi loogika alusel, vaid tõenäosusliku süsteemina, mis loob andmetest üldistusi ja kujundab keelekasutust viisil, mida ei ole alati võimalik täielikult ja arusaadavalt lahti seletada. Seetõttu tuleb läbipaistvust mõista teistmoodi. Rõhk peab liikuma üksikisikute teavitamiselt süsteemi tasandile. See tähendab, et tuleks dokumenteerida, millistest allikatest ja millist tüüpi andmeid kasutatakse, selgitada mudeli eesmärgi ja kasutuspiire ning kirjeldada riske ja nende maandamise võimalusi viisil, mida on võimalik kontrollida ja auditeerida. Samuti peab olema toimiv ja sõltumatu järelevalve. See ei tähenda läbipaistvuse nõude leevendamist, vaid selle kohandamist olukorrale, kus iga üksiku inimese teavitamine ei täida enam tegelikult tema kontrolli tagamise eesmärki.

Analüüs näitab, et anonümiseerimist ja pseudonümiseerimist ei saa käsitada andmete minimaalsuse põhimõtte asendusena, kuna need täidavad erinevat funktsiooni. Anonümiseerimine keskendub andmete mittetuvastatavaks muutmisele pärast nende kogumist, samas kui minimaalsuse põhimõtte eeldab juba eelnevat hinnangut selle kohta, kas ja millises mahus andmete töötlemine on vajalik. Seetõttu ei lahenda anonümiseerimine küsimust andmete kogumise põhjendatusest, vaid võib soodustada ekslikku arusaama, et ulatuslik andmekogumine on lubatav, kui andmed muudetakse hiljem anonüümseks.

LLM-ide kontekstis lisandub risk, et ka puhastatud või anonümiseeritud andmetest võib mudel talletada mustreid, mis võimaldavad hiljem andmete osalist taastootmist või isikut puudutava teabe järeldamist. Sellest tulenevalt ei saa anonümiseerimist pidada lõplikuks ega iseseisvaks kaitsemehhanismiks. Pseudonümiseerimine kujutab endast eeskätt riskide vähendamise vahendit, mis ei vabasta andmetöötlejat GDPR-ist tulenevatest kohustustest. Selle kasutamine eeldab, et isikut tuvastada võimaldav lisateave hoitakse eraldi ning sellele on kehtestatud rangelt piiratud ligipääs, samuti tuleb pidevalt hinnata, kas rakendatud meetmed vähendavad tegelikult tuvastatavuse riski. Nende meetmete roll on toetada kontrollitavust ja riskide vähendamist, mitte õigustada ulatuslikku andmekogumist.

Töös selgus, et andmete minimaalsuse põhimõtte sisuline hindamine võib LLM-ide kontekstis muutuda formaalseks, kui puudub piisav läbipaistvus. Kui ei ole selgelt teada, millistest andmetest mudel koosneb, kuidas neid töödeldakse ja millised riskid sellega kaasnevad, ei ole võimalik hinnata, kas töödeldav andmehulk on eesmärgi suhtes vajalik ega ka seda, kas anonümiseerimine või pseudonümiseerimine vähendab tegelikult tuvastatavuse riski. Ilma selliste kontrollitavate lähteandmeteta jääb hinnang paratamatult oletuslikuks. Kui järelevalveasutustel puudub piisav tehniline pädevus või selge auditiraamistik, piirdub sekkumine sageli üksikjuhtumitele tagantjärele reageerimisega, selle asemel et hinnata ja ennetada süsteemseid riske. Sellest tulenevalt ei saa tehisintellekti andmekaitse tugineda üksnes rikkumistele reageerimisele, vaid eeldab ennetavaid mehhanisme juba süsteemi arendamise etapis. Eelkõige tähendab see selliste dokumenteerimis-, hindamis- ja auditivahendite rakendamist, mis võimaldavad kontrollitavalt tõendada, milliseid andmeid kasutatakse, millises mahus ning kuidas anonümiseerimise või pseudonümiseerimise abil riske tegelikult maandatakse.

Töös leitud ettepanekud võib koondada viide peamisse suunda. Esiteks on vajalik tagada selgem ja süsteemsem koostoime GDPR-i ja tehisintellekti määruse ehk AI Act-i vahel, eelkõige riskipõhise lähenemise järjepideval rakendamisel nii treeningandmete kasutamisel kui ka mudelite väljunditega seotud riskide hindamisel. Teiseks tuleb läbipaistvuse nõude täitmist suunata süsteemsele tasandile, kehtestades andmekogumite dokumenteerimise ja auditeeritavuse standardid. Kolmandaks tuleks LLM-ide puhul andmete minimaalsuse põhimõtet rakendada mõju- ja riskipõhiselt, seostades see selgete mõõdikute ja riskistsenaariumidega, mitte piirdudes üksnes andmemahu abstraktse piiramisega. Neljandaks tuleb tugevdada järelevalvet, kaasates järjepidevalt tehnilist pädevust ning kasutades sõltumatuid auditeid, sertifitseerimist ja standardeid, eriti kõrge riskiga süsteemide puhul. Viiendaks tuleb lõimandmekaitset käsitleda kui sisulist eeltingimust – andmekaitse ei tohiks piirduda vaid dokumentidega, vaid peab mõjutama nii süsteemi ülesehitust kui ka selle kasutamise piire.

Eesti õiguskorras on võimalik ühildada tehisintellekti, sh LLM-ide arendus ja kasutamine GDPR-ist tulenevate andmekaitsepõhimõtetega ilma kehtivat õigust muutmata, kuid see eeldab nende sisulist ja funktsionaalset tõlgendamist. Eesti õigusele omane riskipõhine ja põhjendamiskeskne lähenemine võimaldab käsitada minimaalsust, läbipaistvust ja anonümiseerimist mitte formaalsete nõuetena, vaid kaalutlus- ja riskihindamise osana. LLM-

ide puhul ei hinnata minimaalsust pelgalt andmemahu järgi, vaid seoses riskide ja nende maandamisega, eelkõige andmekaitse mõjuhinnangu kaudu. Praktikas tähendab see näiteks isikut tuvastavate andmete filtreerimist, ligipääsu piiramist ja mudelite testimist võimalike andmeleketega. Oluline ei ole üksnes andmete hulk, vaid see, kas nende kasutamine on põhjendatud ja riskid kontrolli all. Samuti saab Eesti õiguses läbipaistvust käsitada eelkõige süsteemi tasandil, mitte üksnes üksikute andmesubjektide teavitamisena. See tähendab andmeallikate ja -liikide dokumenteerimist, süsteemi auditeeritavust ning järelevalve tegelikku võimaldamist, eriti olukorras, kus individuaalne teavitamine ei ole realistlik. Anonümiseerimine ja pseudonümiseerimine on seejuures vaid riskide vähendamise vahendid, mitte õigustus andmete kogumiseks ega vastutusest vabastamiseks. Need ei asenda hinnangut andmete vajalikkusele ega proportsionaalsusele. Seega ei vähenda LLM-ide areng andmekaitseõiguse tähtsust, kuid näitab, et pelgalt formaalsest vastavusest ei piisa – vajalik on tegelik kontroll, auditeeritavus ja riskide maandamine kogu süsteemi elutsükli vältel.

Käesoleva töö järelduste põhjal ei ole peamine probleem see, et GDPR vajaks ümberkujundamist. Küsimus seisneb pigem selles, kuidas olemasolevaid norme tõlgendatakse ja rakendatakse LLM-ide kontekstis. GDPR on oma olemuselt paindlik ja tehnoloogianeutraalne, võimaldades selle põhimõtteid kohaldada ka keerukate ja andmemahukate süsteemide puhul. Praktilisest vaatenurgast leidis töö, et suurim muutus ei puuduta mitte seda, kui palju andmeid kogutakse, vaid seda, kuidas andmete kasutamist põhjendatakse ja kontrollitakse. Uus teadmine seisneb selles, et LLM-ide puhul ei ole realistlik ega ka sisuliselt piisav keskenduda ainult andmete hulga vähendamisele. Selle asemel tuleks organisatsioonidel juba arendusfaasis süsteemselt läbi mõelda, milliseid andmeliike kasutatakse, milliseid riske see kaasa toob ning kuidas neid riske mõõdetakse ja maandatakse. Praktikas tähendab see näiteks andmekogumite dokumenteerimist, tundlike andmete varajast filtreerimist, mudelite testimist võimalike andmeleketega ning selgete kasutuspiiride kehtestamist. Seetõttu seisneb peamine väljakutse rakenduspraktika arendamises.

Application of the Principles of Data Minimisation and Anonymisation under the General Data Protection Regulation of the European Union to Large Language Models

Abstract

The aim of this master's thesis was to examine to what extent and under which conditions the data protection principles arising from the General Data Protection Regulation (GDPR)—in particular the principle of data minimisation and the requirements related to anonymisation and pseudonymisation—can be reconciled with the development and use practices of artificial intelligence, especially large language models (LLMs). The research problem is framed as a normative dilemma: how to apply the preventive and restrictive logic of data protection law in a context where data processing is large-scale, dynamic, and limited in transparency, and where the value and use cases of data often become apparent only at a later stage.

The central finding of the thesis is that GDPR principles do not lose their validity in the context of artificial intelligence nor become merely formal requirements devoid of practical meaning. However, the analysis demonstrates that their application in the development and use of LLMs is considerably more complex. This complexity stems in particular from the fact that data are collected on a massive scale and from diverse sources, their origin and use are not always fully transparent, and their future uses cannot necessarily be clearly delineated in advance. In such a context, data processing cannot be understood solely as the fulfilment of a predefined purpose, but also encompasses generalisations derived from datasets, the effects of which may manifest later in model outputs and subsequent uses. Thus, the issue is not only one of practical difficulty but also reflects a deeper tension between data protection law and AI development.

With regard to the principle of data minimisation, the thesis concludes that the classical measurability of necessity loses much of its clarity in the context of LLMs. The traditional understanding of minimisation assumes that the necessary amount of data can be determined prior to processing. In the case of LLMs, however, the actual added value of data often becomes apparent only during or after processing. Model performance depends not only on data volume, but also on data quality, provenance, distribution, duplication, filtering, and use cases. Consequently, the substantive application of minimisation cannot be convincingly ensured merely through quantitative restrictions on data volume. Rather, minimisation must be applied in conjunction with necessity, proportionality, and risk assessment, evaluating whether the data

used are justified in relation to the purpose, what risks arise from their processing, and how these risks are mitigated. This approach does not negate minimisation as a normative obligation but adapts its application to the data-intensive and dynamic development logic of LLMs. The key consideration is no longer solely how much data are processed, but what impact the processing has and whether the associated risks are justifiable and controllable.

Regarding transparency, the thesis shows that the conventional approach—based on informing each data subject individually—does not align well with the context of LLM training data. When data are collected from numerous and heterogeneous sources, it is not realistic to identify and inform all affected individuals in the traditional sense. Moreover, explaining the functioning of LLMs is inherently challenging, as they do not operate according to a deterministic input-output logic but as probabilistic systems that generate generalisations and produce language in ways that cannot always be fully and comprehensibly explained. Transparency must therefore be reconceptualised. The emphasis should shift from individual notification to the system level. This entails, in particular, documenting the sources and types of data used, clarifying the model's purposes and limitations, and describing risks and mitigation measures in a manner that is verifiable and auditable. Effective and independent oversight must also be ensured. This does not imply a relaxation of transparency requirements, but rather their adaptation to a context in which individual notification no longer effectively ensures control.

The analysis further demonstrates that anonymisation and pseudonymisation cannot be regarded as substitutes for the principle of data minimisation, as they serve different functions. Anonymisation focuses on rendering data non-identifiable after collection, whereas minimisation requires a prior assessment of whether and to what extent data processing is necessary. Anonymisation therefore does not address the justification for data collection and may instead foster the misconception that extensive data collection is permissible if the data are anonymised at a later stage.

In the context of LLMs, an additional risk arises: even from cleaned or anonymised data, models may retain patterns that enable partial reconstruction of data or inference about individuals. Accordingly, anonymisation cannot be considered a final or standalone safeguard. Pseudonymisation primarily serves as a risk mitigation measure and does not exempt data controllers from GDPR obligations. Its use requires that additional identifying information be

stored separately with strictly limited access, and that continuous assessment be conducted to ensure that the applied measures effectively reduce the risk of identification. These measures are intended to support controllability and risk reduction, not to justify extensive data collection. The thesis shows that the substantive assessment of data minimisation in the context of LLMs may become formalistic in the absence of sufficient transparency. Without clear knowledge of what data constitute the model, how they are processed, and what risks are involved, it is not possible to assess whether the volume of data processed is necessary for the intended purpose, nor whether anonymisation or pseudonymisation effectively reduces identification risks. In the absence of such verifiable baseline information, assessments inevitably remain speculative. If supervisory authorities lack sufficient technical competence or a clear audit framework, intervention tends to be limited to ex post responses to individual cases rather than the evaluation and prevention of systemic risks. Consequently, data protection in AI cannot rely solely on reactive enforcement but requires preventive mechanisms at the system development stage. This includes, in particular, the implementation of documentation, assessment, and auditing tools that enable verifiable demonstration of what data are used, in what volume, and how risks are mitigated through anonymisation or pseudonymisation.

The normative recommendations of the thesis can be summarised in five main directions. First, a clearer and more systematic interaction between the GDPR and the AI Act is required, particularly in the consistent application of a risk-based approach to both training data and model outputs. Second, the fulfilment of transparency requirements should be directed to the system level by establishing standards for dataset documentation and auditability. Third, the principle of minimisation should be applied in a risk- and impact-based manner in the context of LLMs, linking it to verifiable metrics and risk scenarios rather than abstract limitations on data volume. Fourth, the supervisory framework should be strengthened, particularly through the systematic inclusion of technical expertise, independent audits, and certification and standardisation, especially for high-risk systems. Fifth, privacy by design must be treated as a substantive precondition, meaning that data protection must not be confined to ex post documentation but must shape system architecture and usage boundaries.

Within the Estonian legal framework, it is possible to reconcile the development and use of AI, including LLMs, with GDPR data protection principles without amending existing law, provided that these principles are interpreted substantively and functionally. The risk-based and justification-oriented approach characteristic of Estonian law allows minimisation,

transparency, and anonymisation to be treated not as formal requirements but as components of risk assessment and balancing. In the case of LLMs, minimisation is assessed not merely in terms of data volume but in relation to risks and their mitigation, particularly through data protection impact assessments. In practice, this may include filtering personal data, restricting access, and testing models for potential data leakage. The key issue is not the quantity of data but whether their use is justified and risks are under control. Similarly, transparency can be understood primarily at the system level, rather than solely as individual notification. This entails documenting data sources and types, ensuring system auditability, and enabling effective supervision, particularly where individual notification is not feasible. Anonymisation and pseudonymisation are thus tools for risk reduction, not justifications for data collection nor exemptions from responsibility. They do not replace the assessment of necessity and proportionality.

Based on the findings of this thesis, the main issue is not that the GDPR requires reform, but rather how existing rules are interpreted and applied in the context of LLMs. The GDPR is inherently flexible and technology-neutral, allowing its principles to be applied to complex and data-intensive systems. From a practical perspective, the study found that the key shift is not about how much data is collected, but about how the use of data is justified and controlled. The main insight is that, in the context of LLMs, it is neither realistic nor substantively sufficient to focus solely on reducing data volumes. Instead, organizations should already at the development stage systematically consider what types of data are used, what risks this entails, and how those risks are measured and mitigated. In practice, this includes, for example, documenting datasets, filtering sensitive data at an early stage, testing models for potential data leakage, and defining clear usage boundaries. Therefore, the primary challenge lies in developing effective implementation practices.

Kasutatud allikad

Õigusallikad ja valdkondlikud kokkuvõtted

1. Article 29 Data Protection Working Party, *Opinion 05/2014 on Anonymisation Techniques* (WP216), 10.04.2014.
2. Andmekaitse Inspeksioon, *Aastaraamat 2025*, „Olulised kohtuasjad“.
3. European Data Protection Board, *AI Privacy Risks & Mitigations – Large Language Models (LLMs)*, 2025 (Support Pool of Experts Programme).
4. European Data Protection Board, *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*, 17.12.2024.
5. European Data Protection Board, *Statement on generative artificial intelligence and data protection*, 2023.
6. European Data Protection Supervisor, *Opinion 4/2020 on the European Commission's White Paper on Artificial Intelligence – A European approach to excellence and trust*, 19.06.2020.
7. Euroopa Parlamendi ja nõukogu määrus (EL) 2016/679, 27.04.2016, füüsiliste isikute kaitse kohta isikuandmete töötlemisel ja selliste andmete vaba liikumise kohta (isikuandmete kaitse üldmäärus, GDPR).
8. Euroopa Parlamendi ja nõukogu määrus (EL) 2024/1689, 13.06.2024, millega kehtestatakse tehisintellekti käsitlevad ühtlustatud eeskirjad (tehisintellekti määrus, AI Act).

Kirjandus ja muud kasutatud allikad

1. OECD, *Explanatory Memorandum on the Updated OECD Definition of an AI System*, 2023; uuendatud 2024.
2. OECD, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449), 22.05.2019; muudetud 08.11.2023 ja 03.05.2024.
3. European Commission, *Guidelines on the Definition of an Artificial Intelligence System established by Regulation (EU) 2024/1689 (AI Act)*, 06.02.2025.
4. Bender, E. M., *The Social Impact of Large Language Models*, *Communications of the ACM*, 67(1), 2024.
5. Feretakis, G. *et al.*, *GDPR and Large Language Models: Technical and Legal Obstacles*, *Future Internet*, 17(4), 151, 2025.
6. Zhang, R., Li, H.-W., Qian, X.-Y., Jiang, W.-B., Chen, H.-X., *On large language models safety, security, and privacy: A survey*, *Journal of Electronic Science and Technology*, 23(1), 100301, 2025.

7. Zhao, H. *et al.*, *Explainability for Large Language Models: A Survey*, arXiv:2309.01029, 2023.
8. Bommasani, R. *et al.*, *On the Opportunities and Risks of Foundation Models*, arXiv:2108.07258, 2021.
9. Carlini, N. *et al.*, *Extracting Training Data from Large Language Models*, 30th USENIX Security Symposium (USENIX Security '21), 2021.
10. Veale, M., Borgesius, F. Z., *Demystifying the draft EU Artificial Intelligence Act*, *Computer Law Review International*, 22(4), 97–112, 2021.
11. Wachter, S., Mittelstadt, B., Floridi, L., *Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation*, *International Data Privacy Law*, 7(2), 76–99, 2017.
12. Kuner, C., Bygrave, L. A., Docksey, C. (eds), *The EU General Data Protection Regulation (GDPR): A Commentary*, Oxford University Press, 2020.
13. Karu, K., *Tehisintellekti keerukad küsimused*, *Juridica*, 2021/1, lk 43–54.
14. Pärn-Lee, E., *Isikuandmete töötlemise eesmärgipärasus ja selle piirid andmepõhises ühiskonnas*, *Juridica*, 2019/3, lk 175–184.
15. Künnapas, K., *Riskipõhine lähenemine isikuandmete kaitses ja selle piirid*, *Juridica*, 2020/6, lk 401–409.
16. Tupay, P. K., Ebers, M., Juksaar, J., Kohv, K., *Is European data protection toxic for innovative AI? An Estonian perspective*, *Juridica International*, 30, 2021, lk 141–153.
17. Gellert, R., *Data protection: A risk regulation? Between the risk management of everything and the precautionary alternative*, *International Data Privacy Law*, 5(1), 3–19, 2015.
18. Mayer-Schönberger, V., Cukier, K., *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Houghton Mifflin Harcourt, 2013.
19. Sartor, G., Lagioia, F., *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, European Parliament, EPRS, 2020.
20. Biega, A. J., Finck, M., *Reviving purpose limitation and data minimisation in data-driven systems*, *Technology and Regulation*, 2021, lk 44–61.
21. Biega, A. J., Potash, P., Daumé III, H., Diaz, F., Finck, M., *Operationalizing the legal principle of data minimization for personalization*, arXiv:2005.13718, 2020.
22. Shanmugam, D. *et al.*, *Learning to limit data collection via scaling laws: A computational interpretation for the legal principle of data minimization*, FAccT '22, 2022.

23. Dodge, J. *et al.*, *Documenting Large Webtext Corpora: A Case Study on the Colossal Clean Crawled Corpus*, arXiv:2104.08758, 2021.
24. Birhane, A., Prabhu, V. U., Kahembwe, E., *Multimodal datasets: Misogyny, pornography, and malignant stereotypes*, arXiv:2110.01963, 2021.
25. Gebru, T. *et al.*, *Datasheets for Datasets*, arXiv:1803.09010, 2018.
26. Mittelstadt, B., Russell, C., Wachter, S., *Explaining Explanations in AI*, FAT* / FAccT, 2019.
27. Narayanan, A., Shmatikov, V., *Robust De-anonymization of Large Sparse Datasets*, IEEE Symposium on Security and Privacy, 2008.
28. Rocher, L., Hendrickx, J. M., de Montjoye, Y.-A., *Estimating the success of re-identifications in incomplete datasets using generative models*, *Nature Communications*, 10, 3069, 2019.
29. Purtova, N., *The law of everything: Broad concept of personal data and future of EU data protection law*, *Law, Innovation and Technology*, 10(1), 40–81, 2018.
30. Cavoukian, A., *Privacy by Design: The 7 Foundational Principles*, Information and Privacy Commissioner of Ontario, 2011.
31. Iwaya, L. H., Babar, M. A., Rashid, A., *Privacy engineering in the wild: Understanding the practitioners' mindset, organisational culture, and current practices*, 2022.
32. Goldsteen, A. *et al.*, *Data minimization for GDPR compliance in machine learning models*, arXiv:2008.04113, 2020.
33. Ohmann, B., Kollár, G., *Reflections on the data protection compliance of AI systems under the EU AI Act*, *Cogent Social Sciences*, 11(1), 2025.
34. NJORD Estonia, *Is it impossible to be GDPR-compliant in case of AI?*, 2019.
35. Andmekaitse Inspektsioon, *Tähelepanekuid andmekaitse ja tehisintellekti (AI) seostest ja arengutest*, AKI aastaraamat 2024.
36. Andmekaitse Inspektsioon, *Lühiseletused ja seisukohad anonümiseerimise ning kaudse tuvastatavuse kohta*, AKI aastaraamat 2021–2022.
37. European Data Protection Board, *Report on the stakeholder event on anonymisation and pseudonymisation of 12 December 2025*, published February 2026.
38. European Data Protection Board, *Guidelines 01/2025 on Pseudonymisation*, 2025.
39. European Data Protection Board, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, 2020.

40. Euroopa Parlamendi ja nõukogu määrus (EL) 2016/679 2016. *Isikuandmete kaitse üldmäärus (General Data Protection Regulation)*. Euroopa Liidu Teataja L 119, 4.5.2016, lk 1–88.
41. Edwards, L., Veale, M. 2017. *Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for*. *Duke Law & Technology Review*, 16, 18–84.
42. Diakopoulos, N. 2016. *Accountability in algorithmic decision making*. *Communications of the ACM*, 59(2), 56–62. <https://doi.org/10.1145/2844110>
43. Ananny, M., Crawford, K. 2018. *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>

Kohtupraktika

1. Euroopa Liidu Kohus, *Patrick Breyer v Bundesrepublik Deutschland*, kohtuasi C-582/14, ECLI:EU:C:2016:779.
2. Euroopa Liidu Kohus, *CK v Dun & Bradstreet Austria*, kohtuasi C-203/22.
3. Euroopa Liidu Kohus, *European Data Protection Supervisor v Single Resolution Board*, kohtuasi C-413/23 P, ECLI:EU:C:2025:645.
4. Tallinna Halduskohus ja Tallinna Ringkonnakohus, haldusasi nr 3-21-1749 (krediidi- ja maineskooride vaidlus; Kreedix, Register ja Storybook).
5. Tallinna Halduskohus, otsus 9. jaanuar 2025, eraisik vs Andmekaitse Inspeksioon (sotsiaalmeedias avaldatud video ja heli töötlemine ning levitamine). Tallinna Halduskohus, 9.01.2025, haldusasi nr 3-24-2146
6. Tallinna Halduskohus, haldusasi nr 3-23-2208 (maksehäirete avaldamine).

Töö koostamisel kasutati tehisintellekti (ChatGPT, Copilot) idee esialgseks genereerimiseks, keeleliselt sõnakorduste vältimiseks ja töö struktuuri ühtlustamiseks. Kõik lõplikud järeldused ja tõlgendused on autori enda loodud või vastavalt viidatud.

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Katre Kivi ,
(*autori nimi*)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Euroopa liidu isikuandmete üldmääruse isikuandmete minimaalsuse ja anonümiseerimise põhimõtete kohaldamine suurtele keelemudelitele ,

(*lõputöö pealkiri*)

mille juhendaja(d) on Mario Rosentau ,
(*juhendaja nimi*)

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni;

2. annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni;
3. olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile;
4. kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Katre Kivi

26.04.2026