

Käsitsi süvavõltsitud maailm. Postfenomenoloogiline uurimus inimese ja pilditehnika suhte muutustest süvavõltsingute kontekstis

Kevin Rändi, Oliver Laas

Teesid: Artikkel pöörab postfenomenoloogilise pilgu süvavõltsingutele. Postfenomenoloogia on subjektiivsest kogemusest lähtuv mõtteviis, mis uurib, kuidas tehnika kui materiaalsed esemed ja nende kasutusviisid mõjutavad tähendusloomet. Artiklis keskendume süvavõltsingutega kaasnevale inimese ja maailma suhetele, mille keskmes on pilditehnika ja salvestised. Oma käsitluses toome välja, et süvavõltsingute kogemusega kaasneb arusaam digitaalsete salvestiste episteemilistest normidest, mis sarnanevad käsitsi tehtud piltide omadega.

Võtmesõnad: postfenomenoloogia, süvavõltsingud, tehisintellekt, hermeneutiline suhe, pilditehnika

DOI: <https://doi.org/10.7592/methis.v27i34.24692>

Sissejuhatus

Tehisintellekti (TI), täpsemalt masinõppe võimalused kujutiste loomisel ja salvestiste töötlemisel on korraga nii intrigeerivad kui ka problemaatilised. Näiteks ennustatakse, et salvestiste kontrollimatu manipuleerimise tulemusel kaob meie jagatud arusaam faktidest ning tegelikkusest. See on tingitud pilditehnika¹ laialdasest levikust ning olulisest rollist kohtu- ja teaduspraktikas. Salvestised täidavad seal tõendite rolli nendega seotud episteemiliste normide tõttu. Süvavõltsingud (*deepfakes*) – TI abil genereeritud salvestised, milles inimesi kujutatakse tegemas midagi, mida nad tegelikult teinud ei ole – võivad neid episteemilisi norme kahjustada.

Meie hinnangul on senised filosoofilised arutelud süvavõltsingute üle jätnud tähelepanuta kogeva ja tähendusloome protsessides osaleva inimese suhte pilditehnikaga. See on aga oluline osa salvestistega seotud normide kujunemisest süvavõltsingute kontekstis. Fenomenoloogiliselt perspektiivilt lähenedes väidame, et süvavõltsingute puhul on olulised meie muutuvad arusaamad pilditehnikast.

Tehnoloogiliselt reprodutseeritud kujutisi on juba uuritud postfenomenoloogiaks nimetatavas fenomenoloogia ja tehnikafilosoofia harus (Fried ja Rosenberger 2021). Toetutakse arusaamale, et tehnika moodustavad konkreetsetes esemed, mis kasutuse käigus osalevad tähendusloome protsessides ning kujundavad subjekti ja tema kogevavat maailma.

Artikli esimeses osas teemegi täpsemalt juttu postfenomenoloogilise filosoofia kujunemisest ja lähtekohtadest. Postfenomenoloogia põhiliseks hoiakuks on konk-

¹ Pilditehnika hulka kuuluvad salvestusseadmed, salvestised ja töötlusvahendid.

reetsete esemete kasutusviiside ja nende kogemise kirjeldamine, mis avalduvad, kui analüüsime inimese ja tehnika suhteid. Pilditehnika osas peab postfenomenoloogia oluliseks kirjeldada tehnikaga vahendatavat hermeneutilist akti.

Teises osas tutvume lähemalt süvavõltsingute ning nende tekitatud episteemiliste kahjudega seotud filosoofiliste aruteludega. Paljudes neist ei arvestata konkreetsete praktikate ja igapäevasusega, mistõttu eelistamegi postfenomenoloogilist lähenemist.

Kolmandas osas tegeleme postfenomenoloogia enda problemaatikaga. Süvavõltsingutega seotud kogemuste ja praktikate mõtestamiseks on tarvis TI süsteemide ja algoritmide kirjeldust inimese ning tehnika suhetes. Kõnealune kirjeldus viib meid postfenomenoloogia edasiarendusteni, mille kohaselt TI-d ei saa mõista esemena, vaid süsteemina. Samuti väidame, et süvavõltsinguid võimaldaval TI-l on koht meie jagatud arusaamades.

Neljandas osas selgitame süvavõltsingutega seotud algoritmide ja teatavate sotsiotehniliste normide osalust meie tehnikaga vahendatud maailma kogemises. Täpsemalt seisneb süvavõltsingute probleem subjekti jaoks peamiselt suutmatuses eristada autentseid salvestisi võltsingutest. Postfenomenoloogia on kirjeldanud hermeneutilisi strateegiaid pilditehnika vahendatud maailma mõistmiseks (Rosenberger 2008). Kui hermeneutiliselt on oluliseks ka pildilist kogemust moodustav tehniline ese, siis hermeneutiline akt ei piirdu vaid pildil olevaga. Korduv suutmatus eristada võltsinguid autentsetest salvestistest, koos jagatud arusaamaga sellest suutmatusest, muudab meie hermeneutilist suhet pilditehnikaga. Öeldust järeldame, et süvavõltsingute mõjul hakkavad meie digisalvestistega seotud normid sarnanema käsitsi tehtud piltidega seotud normidega. Seda põhjustab süvavõltsingutega kaasnev usaldamatus salvestiste vastu.

Artikli viimases osas näitame, et postfenomenoloogia pakub ka usaldusega seotud eetilisi suuniseid. Neile toetudes pakume välja konkreetseid disainisoovitusi, kuidas süvavõltsingutega haakuvaid episteemilisi probleeme vältida.

1. Tagasi „esemete“ juurde: postfenomenoloogia ja tehnikafilosoofia

20. sajandi esimese poole filosoofia andis olulise panuse tehnikat võimaldavate inimolu tingimuste mõistmiseks. Alates 1980. aastatest rõhutatakse aga, kuidas praktilise elu esemed võivad kanda endas mõningaid ühiskondliku ja kultuurilise elu elemente. Kujunenud hoiak sai hiljem tuntuks kui „empiiriline pööre“ (Achterhuis 2001). Fenomenoloogiline uurimisviis moodustab märkimisväärse osa empiirilise pöörde järgsest tehnikafilosoofiast. Olulisel kohal on Don Ihde (1934–2024) tööd, mis käsitlevad inimese kogemuse intentsionaalset suhet maailmaga, mille keskmes on esemed. Praegu postfenomenoloogiaks kutsutavat haru nimetas Ihde oma teoses „Technics and Praxis“ (1979) esialgu inimese–tehnika suhte fenomenoloogiaks (*phenomenology*

of human-technology relations). Võrreldes näiteks Edmund Husserli, Martin Heideggeri ja Maurice Merleau-Ponty filosoofiaga, on postfenomenoloogias mitmeid erinevusi, kuid ka kokkupuutepunkte.

Ei saa alahinnata Heideggeri olulisust postfenomenoloogia kujunemisest. Muuhulgas saadab postfenomenoloogiat viis, kuidas Heidegger analüüsis tööriistade käepärasust (*Zuhandenheit*) praktikate ja maailmaga suhestumise raames (Ihde 1979, 116–124). Teisalt lükkavad postfenomenoloogid kõrvale Heideggeri romantilise suhtumise inimesesse ning arusaama tehnikast kui millestki abstraktselt (Rosenberger ja Verbeek 2015, 10). Husserlist lähtuvalt saame iga subjektiivse kogemuse puhul rääkida intentsionaalsusest – teadvuse tunnusest, milles kogemus on alati suunatud millegi või kellegi poole maailmas. Ihde tunnistab, et Husserlil oli õigus, kui ta rääkis inimese ja asjade või maailma suhte korreelaadist. Ta selgitab seda skeemiga:

„inimene → maailm“ (Ihde 1979, 6),
 milles „→“ tähistab intentsionaalsust.

Ihde (2016, 96) leiab, et Husserli fenomenoloogias jääb subjekti ja objekti relationiline sidusus kinni „teadvuse keelde“ ja on seetõttu avatud subjektivismi süüdistustele. Lahenduseks on fenomenoloogia ühitamine pragmatismiga: tõsiselt tuleks võtta nii intentsionaalsust kui ka kogemise viise, mis moodustuvad keskkonnas (kehalise) tegevuse vahendusel. Ihde (2016, 131) sõnul mõistsid nii Heidegger kui ka Merleau-Ponty, et instrumentidel on vahendav roll meie kogemuses. Verbeek (2005, 11) toob välja, kuidas tehnika vahendusel (*technological mediation*) moodustuvad nii „mina“ subjektina kui ka „maailm“ minu jaoks.

Nagu Merleau-Ponty, peab ka Ihde oluliseks meie kehastunud suhet konkreetsete esemetega. Kehastumus moodustab põhjaneva suhte maailmaga ja kokkupuutepunkti esemetega. Nii saame rääkida erinevast kogemuse struktuurist. Postfenomenoloogia täiendab eelnevalt mainitud Husserli subjekti ja maailma korreelaadi skeemi „inimene → maailm“. Ihde (1979, 8) nimetab kehastunud suhteks (*embodiment relation*) seda, milles mõni ese on kasutuses kehastunud ning kujundab minu maailma taju:

(inimene–tehnika) → maailm.

Kehastumus on põhjanev, kuid see ei tähenda, et kõik suhted esemete ja maailmaga oleksid sama skeemi alusel mõistetavad. Näiteks meditsiinitöötaja, kes korduvalt teeb ultrahelidiagnostikat, ei pane tähelegi klahvide vajutust ühe käe all ja sondi liigutamist teises käes. Küll aga on talle oluline ekraanil toimuv. Kehatajus muutub tehnika läbipaistvaks, aga näiteks pilditehnika jäädvustusi tuleb „tõlgendada“ ning võib täheldada, et see, kui „läbipaistev“ on meile maailm, sõltub vilumusest instrumendi käsitlemisel (Rosenberger ja Verbeek 2015, 17). Siin joonistub Ihde (1990, 87) töödes teine suhe ja skeem, mida ta nimetab hermeneutiliseks suhteks (*hermeneutic relation*):

inimene → (tehnika–maailm).

Hermeneutilises suhtes saab eristada kahte aspekti: kuidas tehnika vahendab arusaama maailmast (näiteks kellal olev aeg kui meie arusaam ajast); kuidas konkreetseid tõlgendused võimalikuks saavad ja tähendused tõlgendusest ning kasutusest levivad. Ihde (1990, 146) nimetab tehnika kasutuse ja tähenduse variatsioone *multistabiilsuseks* (*multistability*). Selle näiteks on arst, kes hoiab patsiendi ees röntgenipilti, mille tähendus varieerub sõltuvalt patisendi või arsti pilgust. Ihde sõnul on tähenduste variatsioonid osa kõikidest tajudest ja esemete kasutusviisidest. Postfenomenoloogia otsibki selliseid variatsioone, mitte olemust (Ihde 1995). Rosenbergeri (2008) sõnul kasutatakse variatsioonide muutmiseks hermeneutilisi strateegiaid. Näiteks üritab üks meditsiinitöötaja teisele selgeks teha, mida pildil nägema peaks.

Postfenomenoloogiline perspektiiv, et pildiline kogemus moodustub tehnika vahendusel, võib selgitada süvavõltsingute kogemust ning nende episteemilist kahjulikkust. Pildilisel kogemusel on konkreetne (pildi)tehniline tekstuur. See tähendab, et me kogeme pilte ja fotosid oma kehastunud tajus esemelistena – paberile trükituna, raamistatuna, pildialbumis, galeriiseinal. Digitaalsete salvestiste esemelisuse puhul tulevad mängu ekraanid, veebirakendused, kasutajaliideste elemendid ning jagamise, kommenteerimise jt funktsioonid. Need kõik on osa meie kogemusest ja hermeneutilisest suhtest tehnikaga. Pildiline kogemus on alati juba esemete poolt töödeldud. Tehnikaga kaasnevad mitmed kasutus- ja tähendusvariatsioonid. Kui soovime uurida hermeneutilist suhet – kuidas me tehnikast midagi maailma kohta välja „loeme“ –, tuleb arvestada, et pildilise kogemuse multistabiilsed variatsioonid sõltuvad juba kehalisest paigutusest ja esemetest meie ümber.

Salvestiste multistabiilsusele ning tehnilisele tekstuurile keskendumise vajadus tuleneb algoritmide interaktiivsusest ja võimest kasutajaid suunata. Digitaalsete esemete (Wiltse 2014) ning algoritmide võime kujundada meie kogemust ja käitumist (Wellner 2020) tekitab vajaduse ümber mõtestada mõningad Ihde intentsionaalse inimese–tehnika suhte skeemid. Selle juurde pöördume tagasi artikli 4. osas.

2. Episteemilise kahju käsitluste problemaatika

Levinud on nii visuaalsed kui ka helilised süvavõltsingud. Keskendumise siinkohal visuaalsetele süvavõltsingutele, sest heli kogemus (kuigi käsitletav hermeneutiliselt inimese–tehnika suhtena) erineb nägemisest piisavalt, et selle lisamine ületaks käesoleva artikli piirid.

Süvavõltsinguid on kasutatud ka episteemilise kahju tekitamiseks. Selmet alustada „episteemilise kahju“ määratlusest, alustame näidetega, sest määratlused on seotud filosoofiliste episteemilise kahju käsitlustega. Mõningad elulised näited episteemilise kahju põhjustamisest:

- ajakirjaniku maine kahjustamine temast süvavõltsitud pornograafia levitamise kaudu (Ayyub 2018);
- poliitiku poolt valijatele oma keelelistest oskustest eksitava mulje jätmine (Christopher 2020);
- Belgia peaministri (Galindo 2020) ja Vene ning Ukraina presidentide suu läbi valeinfo levitamine sotsiaalmeedias (Baig 2022);
- autentsete salvestiste süvavõltsingutena käsitlemine poliitilisest skandaalist pääsemiseks (Schick 2020).

Viimastel aastatel on ilmunud mitmeid filosoofilisi seletusi süvavõltsingute episteemilisest kahjulikkusest. Uurima on hakatud ka süvavõltsingutega kaasneda võivat düstooptiat – infokalüpsis või episteemiline apokalüpsis –, mis puudutab kogu ühiskondlikku elu ja reaalistaju (Habgood-Coote 2023). Järgnevad vastuargumendid aga puudutavad argipraktikate ja hoiakute rolli ning nendega seonduvaid puudujääke süvavõltsingute episteemilise kahjulikkuse käsitlustes. Samuti suunavad nad meid süvavõltsingute postfenomenoloogilise käsitluse poole ning toovad esile selle vajalikkuse.

Rini (2020) leiab, et salvestised moodustavad nn episteemilise piirde, sest me toetume neile tunnistuste kontrollimisel. Ilmekaks näiteks on kohtupraktika, kus turvakaamerate salvestised tõenditena on pealtnägijate tunnistustest kaalukamad. Salvestised täidavad episteemilise piirde funktsiooni nendega seotud episteemiliste normide abil. Süvavõltsingud on episteemiliselt kahjulikud, sest nad õõnestavad neid norme. Fallis (2020) leiab, et see käsitlus on poolik, sest ei seleta, kuidas süvavõltsingud salvestistega seotud episteemilisi norme õõnestavad. Samuti ei ole selge, miks me ei saaks salvestistele toetumise asemel tunnistusi kontrollida teiste tõendite või allika usaldusväärsuse alusel.

Fallise (2020) sõnul on süvavõltsingud episteemiliselt kahjulikud, sest nad kahan-davad salvestiste poolt edastatava informatsiooni hulka. Vaadakes salvestisi neil kujutatud olukordade kohta informatsiooni edastavate signaalidena. Tõese d p o s i t i i v s e d signaalid kujutavad salvestise tegemisel aset leidnud olukordi, v a l e p o s i t i i v s e d aga mitte. Süvavõltsingud on episteemiliselt kahjulikud, sest nad suurenda-vad salvestiste populatsioonis valepositiivsete kohtamise tõenäosust ja kahandavad sellega salvestiste usaldusväärsust tõenditena.

Sellel käsitlusel on mitmeid puudusi. Kui tõeste positiivsete hulk kasvab koos valepositiivsete hulgaga, mis paistab parasjagu juhtuvat, siis ei suurenda süvavõltsingud, vähemalt praegu, valepositiivsete sagedust salvestiste populatsioonis (Carlson 2021). Samuti ei ole Fallise seletus psühholoogiliselt usutav, sest salvestisega kokku puutuva inimese kõhklused ei tulene tõeste ning valepositiivsete esinemissagedusi puudutavatest teadmistest, vaid sellest, et ta ei suuda neil vahet teha (Laas 2023).

Carlson (2021) väidab, et süvavõltsingud on episteemiliselt kahjulikud, sest loovad lõhe olukordade ja salvestiste kui nende kohta käivate tõendite vahele. See lõhe viib skeptitsismini selle osas, kas salvestised tõenditena saavad õigustada uskumusi neil kujutatud olukordade kohta. Carlsoni arvates on tegemist kartesiaanliku deemoni sarnase olukorraga, sest ka deemoni hüpotees tekitab lõhe meeleanndmete kui tõendite ning välismaailma puudutavate teadmiste vahele. Erinevalt kartesiaanlikust skeptitsismist esitavad süvavõltsingud meie teadmistele praktilise väljakutse, sest suutmatus võltsinguid autentsetest salvestistest eristada mõjutab meie argipraktikaid.

Ka sellel käsitlusel on omad puudused. Süvavõltsingud mõjutavad meie argipraktikaid ainult siis, kui nad muudavad meie hermeneutilist suhet salvestistega ehk seda, kuidas me neid tõlgendame, mis omakorda tähendab seda, et skeptitsism on mingisuguse sügavama põhjuse tagajärg, mitte ise süvavõltsingute kahjulikkust seletav põhjus (Laas 2023).

Matthews (2022) kirjutab, et süvavõltsingud on episteemiliselt kahjulikud, sest soosivad intellektuaalset künismi – negatiivset hoiakut internetivideote kui teadmiste allikate ning tõendite suhtes. Selle tulemusel kujuneb välja küüniline tundelaad ehk harjumus internetivideosid ebausaldusväärsetena näha. Intellektuaalne künism on episteemiline pahe, väidab Matthews, sest selle mõjul takistab küünik informatsiooni levikut, pidades usaldusväärseid videoid ebausaldusväärseteks.

See seletus ajab segamini põhjuse ja tagajärje: intellektuaalne künism on hoiaku muutuse tagajärg, mitte selle põhjus. Kasutaja küüniline tundelaad kujuneb välja usaldamatusest internetivideote vastu, mis omakorda tuleneb korduvatest eksimustest autentsete videote võltsingutest eristamisel (Laas 2023). Selline küünilise tundelaadi kujunemine sarnaneb teiste inimeste suhtes küüniliseks muutumise protsessiga, mis algab usaldust õõnestavatest kogemustest.

3. ...Aga kui tehisintellekt pole „ese“

Ka postfenomenoloogial on süvavõltsingute käsitlemisel omad piirangud. Liiga esemetekeskne lähenemine jätab varju TI-le iseloomulikud tunnused, laiema tähenduste ja kultuurisõltelise tausta (Coeckelbergh 2022). TI-d on raske taandada pelgalt esemele (Crawford 2021). Coeckelbergh osutab tööikadele, kus TI astub erisugustesse sotsiaalsetesse rollidesse (2020) ning kujundab meie narratiive TI tulevikust (2021). Isegi kui postfenomenoloogia kohaselt konkreetne TI ilmneb meie jaoks selle teisena – justkui „kunstnik“ või „kaasautor“ –, jäävad nende mitmed tähenduslikud rollid märkamata, kui ei arvestata, kuidas strukturalistlikult, kultuurisõltelisel ja keelelisel tasandil arusaamad TI-st kujunevad (Coeckelbergh 2022).

Coeckelberghi (2022, 151) sõnul tuleks filosoofidel uurida TI transtsendentaalseid tingimusi või sotsiaalset mõõdet nõnda, et arusaam tehnikast jääks konkreetseks. Üks

võimalus selleks on läheneda süsteemikeskselt, kus tähelepanu on nii TI tehnilistel kui ka sotsiokultuuriliste kasutuspraktikatega seotud aspektidel. Alustame eristusest digitaalsete fotograafia ja traditsioonilise analoogfotograafia vahel ning analüüsime nende poolt vahendatud maailma kogemise viiside episteemilisi erinevusi.

Fotode vahendusel maailmas leiduvate olukordade kogemise taustal on üldlevinud teadmised traditsioonilisest fotograafiast kui tehnoloogiast. „Tehnoloogiat“ mõistame siinkohal sotsiotehnilise süsteemina, mille moodustavad:

1. tehnilised esemed,
2. inimtoimijad ja
3. nende poolt järgitavad institutsioonid või reeglid (van de Poel 2020, 391).²

Traditsiooniline fotograafia kui tehnoloogia on samuti sotsiotehniline süsteem, mis koosneb kaameratest, nende (osade) valmistajatest, kasutajatest, fotode levitamise viisidest, ühiskondlikust tööjaotusest jne.³ Need asjad koos muudavad fotograafia kui sotsiokultuurilise visuaalse kommunikatsiooni praktika võimalikuks.⁴

TI süsteemid on sotsiotehnilised süsteemid, milles inimtoimijaid täiendavad või asendavad autonoomsed ja interaktiivsed tehnilikud toimijad – masinõppe mudelid, (jutu)robotid – ja institutsioone asendavad või täiendavad tehnilised normid – põhjuslik-füüsikalised reeglid (algoritmid), mis suunavad tehnilike toimijate käitumist ning on osa toimija tarkvarast või TI süsteemi disainist (van de Poel 2020). Tehnilised normid ei piirdu TI süsteemidega. Nad võivad kodifitseerida episteemilisi, juriidilisi ja moraalseid norme tehnikatesse, mõjutades nende disaini viisil, mis suunab kasutajaid moraalselt ning juriidiliselt soovitud käitumisviiside poole ja eemale soovimatutest käitumisviisidest (van den Berg ja Keymolen 2017).

Süvavõltsinguid võib pidada TI süsteemiks eespool toodud tähenduses. Nad koosnevad riistvarast (kaamerad, arvutid), tarkvarast (kaasa arvatud masinõppe mudelid) süvavõltsingute loomiseks ja nende levitamise kanalitest. Selle süsteemi osade käitumist suunavad sotsiaalsed ning tehnilised normid.

Traditsioonilistel kaameratel ja neid ümbritseval sotsiotehnilisel süsteemil on teatud omadused, mida teavad kaudselt ka pelgalt fotosid vaatavad inimesed ning mis mõjutavad nii nende vaatamiskogemust, fotode poolt vahendatud suhet maailmaga kui ka fotode tõlgendamise viisi.

2 Alternatiivsete käsitluste seast eelistame seda, sest see on hõlpsalt laiendatav süvavõltsingutele kui TI tehnoloogiale.

3 Ühiskondliku tööjaotuse all peame silmas, et nii traditsiooniliste kaamerate käsitsemise kui ka filmi ilmutamise tehnilised oskused, nagu ka kaamerate valmistamise ning parandamise tehnilised oskused on kogukonnas jaotunud sellisel, et saame eristada spetsialiste tavakasutajatest, professionaalseid fotograafe hobifotograafidest.

4 Tehniliste esemete määratlus erineb sotsiotehniliste süsteemide määratlusest selle poolest, et viimane hõlmab eri gruppide ja institutsioonide kohati vastakaid huve ning arusaamu mingi tehnilise eseme sihipärasest kasutamisest. Tehnilised esemed saavad oma sihid ja kasutusviisid sotsiotehnilistest süsteemidest, millesse nad kuuluvad ning milles parasjagu mõne huvigrupi arusaamad prevaleerivaks osutuvad.

Esimene asjakohane omadus on, et traditsiooniline fotograafia on informatsiooni säilitav põhjuslik süsteem, mis piirab inimeste võimalusi fotode tegemise etappides neisse uut informatsiooni sisestada (Hopkins 2012). Seetõttu pakuvad traditsioonilised fotod kaudset faktilist pildilist kogemust neil kujutatud asjaoludest.⁵

Teist asjakohast omadust kirjeldab Kendall Waltoni (1984) läbipaistvuse tees, mille kohaselt fotod on läbipaistvad selles mõttes, et x -i kujutava foto vaatamisel me näeme kaudselt ja vahendatult x -i.⁶ Näiteks praegust Eesti presidenti kujutava foto vaatamisel me näeme artikli kirjutamise ajal vahendatult Alar Karist.

Need kaks omadust on osa ühisest teadmisesest digitaalse fotograafia eelses kultuuris, millele pädevad vaatajad toetuvad. Traditsiooniliste fotode poolt pakutav kaudne faktiline pildiline kogemus toetab tavatingimustes vaataja vaikivat doksastilist reaktsiooni (*default doxastic response*), mille kohaselt ta usub, et P , kui fotol on kujutatud propositsiooni P kinnitav olukord (Cavedon-Taylor 2013, 295).⁷ Näiteks tavatingimustes Alar Karist kujutavat traditsioonilist fotot vaatav inimene kogeb, et ta näeb selle kaudu Karist, mis omakorda tekitab vaikiva doksastilise reaktsiooni, mille kohaselt ta usub fotol kujutatut (näiteks, et Karis kandis pildistamise ajal tumesinist pintsakut).

Hopkins (2012) leiab, et digifotograafial ei ole traditsioonilise fotograafia kahte omadust. Põhjus on selles, et digifotodesse uue informatsiooni lisamine on oluliselt lihtsam. Seda kinnitavad nii fototöötlustarkvara pakutavad võimalused kui ka nutitelefoni kaamerarakendustesse programmeeritud filtrid, mis kahandavad digifotodel kujutatu muutmiseks vajalike oskuste taset. Hopkins järeldab, et digifotod ei paku kaudset faktilist pildilist kogemust, mistõttu nad ei ole Waltoni mõttes läbipaistvad ega tohiks esile kutsuda eespool kirjeldatud vaikivat doksastilist reaktsiooni.

Leiame, et Hopkins püüab luua aredat eristust, mida praktikas ei esine. Digifotod võivad samuti esile kutsuda vaikiva doksastilise reaktsiooni, mille kohaselt me usume, et asjaolud on või olid nii, nagu digifoto neid kujutab. Seda kinnitavad inimeste reaktsioonid digifotodele ja nende kasutamine asitõenditena kohtus.⁸ Seega leiame, et siin on tegemist pigem määra kui liigi erinevusega.

5 Faktiline pildiline kogemus on kogemus, milles me „haarame, mida pilt kujutab“ ja milles, „kui pildil nähakse, et p , siis p “ (Hopkins 2012, 2, 3). Nimetame seda kogemust siin kaudseks, sest soovime seda seostada traditsioonilise fotograafia teise asjakohase omadusega.

6 Algselt esteetilist teesi on hiljem hakatud kasutama fotode episteemilise staatuse ja nendega seotud episteemiliste normide seletamiseks (Cavedon-Taylor 2013, 284 m. 2). Väidame, et teesi sisu võib tavavaatajale intuiitiivselt tuttav olla varasemate vaatamiskogemuste ning fotograafiaga seotud kultuuriliste ja episteemiliste normide põhjal.

7 Kuigi räägime siin fotodest, laieneb analüüs tavatingimustes ka (traditsioonilistele) analoogvideotele.

8 Hopkins võiks vastu väita, et meie digifotosid hõlmav praktika on episteemiliselt ekslik, kuid leiame, et tema eristus võib küll kehtida filosoofia ideaalses maailmas, aga mitte tegelikkuses, sest massilise eksimise omistamine nii suurele hulgatele inimestele nõuaks Hopkinsi poolt esitatutest tugevamaid tõendeid.

Erinevalt traditsioonilistest kaameratest, mis olid mehaanilis-keemilised süsteemid ja kujundasid sellistena meie pildilist kogemust, on nüüdisaegsed digikaamerad valdavalt TI süsteemid, milles meie pildilist kogemust kujundavad ka tehnilised normid. Nii nutitelefonide kui ka digikaamerate tehtud pildid on algoritmiliselt töödeldud, sest nende sensorite töö on tarkvara poolt kontrollitud. Näiteks hämarates tingimustes tehtud digifotode puhul korrigeerivad uuemad nutitelefonid automaatselt pildi kvaliteeti lähtuvalt tarkvara osaks olevate masinõppe algoritmide ennustustest selle kohta, mida pildistaja neis (või ideaalsetes) tingimustes näeks.

4. Hermeneutiline suhe tehnikaga ja salvestustehnika „intentsionaalsus“

Tuleme tagasi üksiksubjekti, inimese ja tehnika suhete ning intentsionaalsuse juurde, et tuvastada, kuidas eespool öeldu avaldub konkreetsetes inimese–tehnika suhetes. Argikogemuses ilmneb süvavõltsingute probleem ennekõike individuaalsel tasandil, kus üksikisik ei suuda autentseid salvestisi võltsingutest eristada.

Seoses algoritmidega saab rääkida tehnikale omasest „intentsionaalsusest“, mis lähtub tehnika valitud reaktsioonist ning võimest „ilmutada uusi aspekte reaalsusest“ (Wellner 2020). Ihde (1979, 77–78) ja Verbeek (2008, 392) toovad näiteks diktofoni, millega soovime jäädvustada meile tähenduslikku heli (nt õppejõu juttu), kuid diktofon salvestab sama valjult taustamüra (nt kedagi tagareas kõhinas). Verbeek (2008) leiab, et kunstis ja mujal kasutatakse ära seda, et tehnika „märkab“ muud või lausa loob seda, mida inimese enda tajud ei pruugi võimaldada. Seda saab kirjeldada inimese–tehnika suhtena, milles „esemed ise mängivad kesket rolli inimese intentsionaalsusega koostöös“ (Verbeek 2008, 392). Seda nimetatakse postfenomenoloogias liitsus-suhteks (*composite relation*):

inimene → (tehnika → maailm) (Verbeek 2008, 393).

Öeldu valguses pakume välja, et meie digifotode ja -videote poolt vahendatud suhet tegelikkusega saab iseloomustada järgmiselt:

inimene → (fototehnoloogia – tehnilised normid → maailm).⁹

Meie skeemis on sulud samal põhjusel mis ülemise, hermeneutilise suhtega sarnase liitsussuhte puhulgi: tehnika ilmutab maailma, kuid teeb seda ka teisele kuuluva intentsionaalsusega. Skeemi puhul on oluline, et kui teisele kuuluva intentsionaalsusega segunevad normid, fototehnoloogia ja algoritmid, siis on seda kõike tajuval inimesel raske eristada, kelle või mille intentsionaalsus see on.¹⁰

⁹ Siin käivad tehniliste normide alla ka algoritmid.

¹⁰ Normide rõhutamine intentsionaalsusest rääkimisel on oluline, sest kuigi Verbeek ja teised on valmis heaks kiitma intentsionaalsust tehnika tunnuseks, võib tekkida arusaamatus, kuna fenomenoloogias on intentsionaalsus teadvuse tunnus. TI puhul võib tehnilisest intentsionaalsusest muidugi rääkida, aga rõhutades, et selles justkui on intentsionaalsus meie jaoks.

Verbeek (2008) näitlikustas liitsussuhet kaameraga kunstiteoste loomise abil. Kunstnik adub, et kaamerasilm võimaldaks jäädvustada midagi, mida inimsilm tajuda ei suuda. Sellistel puhkudel suudab vaataja eristada oma tajuviisi jäädvustuse omast, mis kutsub teda viimast tõlgendama. Süvavõltsingute puhul on aga iseloomulik olukord, kus TI süsteem salvestises jääb märkamatuks ja – postfenomenoloogilises tähenduses –, selline salvestis on hermeneutiliselt sama läbipaistev kui autentne „reaalse“ sündmuse jäädvustus.

Mõned variatsioonid on aga veel võimalikud. Võtame näiteks kellegi, kes üritab diktofoni „üles seada“, et auditoorset loengut salvestada. Ta püüab tähelepanu allikale võimalikult lähedale minna, sest teab, et diktofon võib subjekti jaoks üles korjata taustamüra. Pädev lindistuste tõlgendaja oskab selles olukorras konkreetse tehnikaga ennetavalt suhestuda. Ühest küljest on see kehastunud suhe, kuna ta liigutab ruumis diktofoni justkui oma kunstliku ja kehast väljaulatuvat „kõrva“. Kuid võimalik on ka hermeneutiline suhe: konkreetne salvestustehnika ei ole ainult keha laiendus, vaid pädeva kasutuse moodustab ideaaltingimustel see, kuidas kasutaja tõlgendab salvestusinstrumenti ennast – selle funktsionaalsust ja iseloomulikkust (tehnilist „intentsionaalsust“), et ümbritsevast tähenduslikku infot saada.

Seega leidub kasutamise variatsioone, inimese ja tehnika suhete võimalusi, mille põhjal saame öelda, et moodustuvad suhted tehnikaga, mille analüüsimisel ilmnevad eespool mainitud süsteemsed faktorid. Saame kirjeldada, kuidas võib muutuda hermeneutiline suhe digisalvestistega. Neid muutusi aitab mõista Ihde (1990, 75) tähelepanek, et tehnika ei muutu „minu“ ja „maailma“ vahel kunagi täiesti „läbipaistvaks“ – nt prillid ei asenda silma, kuna võivad hõõruda ja tuleb puhastamiseks eemaldada. Teisisõnu on tehnikal materiaalsed ja funktsionaalsed piirangud ning määravaks osutuvad ka mitmesugused normid, mis nõuavad meie tähelepanu. Öeldu põhjal saame rääkida, a) kuidas me tõlgendame maailma tehnika kaudu ja muudame tõlgendusi (Rosenberger 2008); ning b) kuidas me tõlgendusi muudame vastavalt tehnika iseloomule maailma vahendada.

Muutuste paremaks mõistmiseks võrdleme digifotode poolt vahendatud suhet maailmaga käsitsi tehtud kujutiste, joonistuste ja maalide vahendatuga, sest ka siin sõltub maailma vahendava kujutise sisu teise osalise, käsitsi tehtud kujutise autori kavatsustest ning tõlgendustest. Alustame taaskord käsitsi tehtud kujutiste kui tehnoloogia omaduste kirjeldamisega.

Hopkins (2012, 1–2) leiab, et oma hõlpsa manipuleeritavuse tõttu sarnanevad digifotod pigem käsitsi tehtud piltide kui traditsiooniliste fotodega. Ta lisab, et käsitsi tehtud pildid võivad kajastada fakte. Illustratsioon tehnilises käsiraamatus, õlivärvidega tehtud portree või reporteri visand kohtusaalist võivad kõik täpselt tabada kuju-

tatud stseene. Kui nad seda teevad, siis me näeme neis piltides seda, kuidas asjad tegelikult on (või olid).

Aga faktilisus nõuab enamat kui faktide tabamist, see nõuab nende paratamatut tabamist. Käsitsi tehtud pildid ei tee seda mitte kunagi, sest isegi kui nad on täpsed, siis on nad seda seetõttu, et nad kajastavad, kuidas keegi – kes iganes pildi tegi – asjaolusid tajus. Selle tulemusel on käsitsi tehtud piltide puhul alati vea oht, hoolimata sellest, kas neis tegelikult vigu esineb. (Hopkins 2012, 6–7)

Teiste sõnadega on käsitsi tehtud pildid täpsed siis, kui nad kujutavad seda, kuidas nende autorid pildi tegemise hetkel kujutatud asjaolusid õigesti ja täpselt tajusid. Inimautori intentsionaalsuse ja tõlgenduste kujundav ning vahendav roll käsitsi tehtud piltide loomisel on usutavalt ühine teadmine meie kultuuris. Skemaatilisel saame käsitsi tehtud piltide poolt vahendatud suhet maailmaga kujutada järgmiselt:

inimene → (inimene–pilditehnika → maailm).¹¹

Sellel suhtel on olulised tagajärjed nii meie teadmiste kui ka pildilise kogemuse jaoks. Kui faktilisus eeldab paratamatust, siis meie pildiline kogemus käsitsi tehtud piltidest ei saa olla faktiline. Kuna me ei saa käsitsi tehtud pildi enda põhjal otsustada, kas see kujutab fakte tõepäraselt või mitte, siis saab pildi autorist selle faktilisust määrav faktor.

Näiteks võõras linnas teed küsides me usume vastuseks saadud juhiseid, kui me peame nende autorit usaldusväärseks. Oletame, et keelebarjääri tõttu ei ole verbaalsed juhised arusaadavad. Kui meie kaasvestleja saab sellest hoolimata aru, mida me küsime ja kuhu soovime jõuda ning tema joonistatud juhiste järgimine viib meid soovitud sihtkohta, siis täitis joonistus sama kommunikatiivset funktsiooni mis verbaalsed juhisedki. Nagu verbaalsete juhiste puhul, otsustame ka joonistatud juhiste täpsuse üle autori usaldusväärseuse põhjal.

Seda, et verbaalne tunnistus ja joonistus võivad kanda sarnast informatsiooni ning me hindame mõlema faktilisuse üle autori usaldusväärseuse põhjal, kinnitab veel järgmine näide.¹² Oletame, et meist kumbki ei ole kunagi Alar Karist näinud ning foto- ega videosalvestisi pole samuti käepärast. Meil on aga kaks joonistust, esimene tuntud karikaturisti ja teine lugupeetud portretisti poolt. Kummast joonistusest peaksime lähtuma, kui soovime tuvastada, kas restoranist väljuv mees on Alar Karis? Tundub, et kindlam oleks toetuda portretisti joonistusele, sest temal on parem maine loodud kujutiste täpsuse ja tõepärasuse osas.

11 Siinkohal peame pilditehnika all silmas käsitsi tehtud pildi valmistamiseks kasutatud vahendeid: estampgraafikat, õlimaali, pliiatsijoonistust jne.

12 Episteemilist sarnasust tunnistuse ja joonistuste vahel kinnitavad ka ajaloolised näited joonistatud kaartidest ning tagaotsitavatest.

Õeldust tuleneb, et käsitsi tehtud piltide puhul sõltub meie kogemuse faktilisus nende päritolust ehk inimesest eespool toodud suhte „inimene–pilditehnika“ koosluses. Kui Hopkinsil on õigus, et digifotod sarnanevad nii kogemuslikult kui ka episteemiliselt pigem käsitsi tehtud piltidega, siis sõltub meie kogemuse faktilisus ka nende puhul traditsioonilise fotograafiaga võrreldes enam digifoto päritolust.

Süvavõltsingud ei too selles osas kaasa sisulisi muutusi. Pigem kiirendavad nad veelgi digifotograafia poolt algatatud protsesse, sest teevad senisest lihtsamaks inimeste tõlgenduste fotodesse ning videotesse sisestamise. Süvavõltsingud muudavad fotod ja videod episteemiliselt ning kogemuslikult käsitsi tehtud piltide ja animatsioonide sarnaseks (Laas 2023). Meie suhet maailmaga, süvavõltsitud kujutise poolt vahendatuna, võiks üldistatult kujutada nii:

inimene → (TI süsteem → maailm).

Selles skeemis hõlmab TI süsteem tehnoloogilisi esemeid (foto- ja videokaamerad), inim- ja tehismoimijaid ning tehnilisi norme. Sellise koosluse puhul saame rääkida erisugustest intentsionaalsustest ning nende vastasmõjust. See suhe on sarnane käsitsi tehtud piltide poolt vahendatud suhtega, kuid inimautori poolseid, fotol või videol kujutatut kujundavaid tõlgendusi mõjutavad siin autonoomsete tehnilike toimijate (masinõppe mudelite) „tõlgendused“.

Sarnaselt käsitsi tehtud pildiga ei ole ka süvavõltsitud kujutise pildiline kogemus faktiline. Erinevalt käsitsi tehtud pildist on süvavõltsitud kujutise pildilise kogemuse osaks teatud pettus või semiootiline mimikri. Kvaliteetselt süvavõltsitud kujutis kutsub esile (digi)fotoodega seotud vaikiva doksaatilise reaktsiooni, sest jätab mulje, et oleme selle vahendusel suhtes „inimene → (fototehnoloogia – tehnilised normid → maailm)“, kuigi tegelikult oleme hoopis skeemiga „inimene → (TI süsteem → maailm)“ kirjeldatud suhtes, mis eeldaks hoopis käsitsi tehtud piltidele omast vaikivat doksaatilist reaktsiooni. Teiste sõnadega, süvavõltsitud kujutis tekitab kogemuse, mille varal me usume kujutisel kujutatu põhjal, et *P*, kuigi peaksime hoopis uskumuste moodustamisest hoiduma ning kujutise päritolu tuvastama, et selle põhjal kujutise tõepärasuse üle otsustada.

5. Usaldus, süvavõltsingud ja postfenomenoloogiline eetika

Lisaks suhete kirjeldusele peaks postfenomenoloogiline uurimus aitama kujundada ka tehnika suhtes normatiivseid seisukohti (Verbeek 2019). Järgnevalt proovime seda teha seoses süvavõltsingutega.

Tuleks arvestada võimalusega, et võimetus autentseid (digi)fotosid ja -videosid süvavõltsingutest eristada, koos korduvate eksimustega, viib usaldamatuseni sal-

vestustehnoloogia vastu (Laas 2023).¹³ Usaldus ja usaldamatus on tõlgenduslikud hoiakud, mis mõjutavad seda, kuidas me oma kogemusi ning käepäraseid tõendeid tõlgendame (Govier 1997, 34, 35). Seetõttu võib pelk teadmine süvavõltsingute olemasolust ja võimalikkusest viia salvestustehnoloogia vastu tuntava usalduse kaahanemise või kadumiseni. Sõltumata sellest, kas usaldamatus tuleneb teadmisest või isiklikust kogemusest, võib see nakatada kõiki meie salvestistega seotud pildilisi kogemusi.

Meie arutelu põhjal on võimalik, et usaldamatus salvestustehnoloogia vastu toob kaasa hoiakulise muutuse. Selle tulemusel hakkab salvestiste kogemus sarnanema käsitsi tehtud kujutiste kogemusele. Kui käsitsi tehtud kujutiste usutavus on seotud nende allika usaldusväärusega, siis hoiakulise muutuse tagajärjel hakkab ka salvestiste usutavus sõltuma nende allikate usaldusväärusest. Kuidas sellises olukorras kindlaks teha tehniliste salvestise allika usaldusväärus?

Üheks lahenduseks on salvestise päritolu ja allika usaldusvääruse autentimissüsteemi loomine. Selline süsteem peaks aga ise usaldusväärne olema. Laas (2023) pakub välja kolm disainiomadust, mis sellisel süsteemil peaksid olema, et see saaks kasutajatele oma usaldusväärust tõendada:

1. Usaldusväärne allikate autentimise süsteem peaks võimaldama kasutajatel korduvalt andmeid autentida ja nende allikaid tuvastada.
2. Usaldusväärne autentimise süsteem võiks olla detsentraliseeritud.
3. Usaldusväärne autentimise süsteem võiks olla avatud lähtekoodiga.

Kasutaja perspektiivilt võib neid omadusi käsitada heuristikatena, mis aitavad tehnoloogiliste lahenduste vahel valikut langetada.

Usaldusväärne autentimise süsteem peaks võimaldama salvestise päritolu ja selle allika usaldusvääruse korduvat tuvastamist. Selleks on vaja kindlat kasutuse ahelat, mis näitaks, millal ja kelle poolt mingisugune kujutis esmakordselt interneti ning meediasse toodi.

Tsentraliseeritud süsteemide peamiseks nõrkuseks on teenusepakkuja infrastruktuur, mille tõrke korral ei saaks kasutajad enam selle abil süvavõltsinguid autentsetest kujutistest eristada. Detsentraliseeritud süsteemidel, nagu P2P võrgud, selliseid probleeme ei ole, sest teenuse pakkumiseks vajalikke funktsioone täidavad paljud instantsid ning süsteem tervikuna ei ole ühe institutsiooni kontrolli all. See on positiivne, sest „informatsiooni filtreeriva võimu koondamine üksikute, kasumist ajendatud tehnoloogiahiidude kätte on demokraatialle ohtlik“ (Ford 2021, 280). Oht on seda suurem, mida

¹³ Vastu saaks väita, et usaldus on hoiak, mida ei saa võtta tehnoloogia suhtes. Kuna Laas (2023) põhjendab, miks levinud arutlused tehnoloogia usaldamise vastu ei ole veenvad, siis ei hakka me siinkohal neid vastulauseid korradama, vaid soovitame lugejal seal toodud põhjendusi iseseisvalt hinnata. Siin ja edaspidi eeldame, et usaldus tehnoloogia vastu ei ole aprioorsetelt võimatu.

olulisemat rolli mängib säärane süsteem sisuliselt meie kõigi jagatud tegelikkuse loomises osalevate kujutiste valideerimises.¹⁴

Viimaks peaks usaldusväärne autentimise süsteem olema avatud lähtekoodiga, sest siis oleksid selle usaldusväärset näitavad tõendid avalikult kättesaadavad ja kõigi huviliste poolt kontrollitavad. Isegi kui kõik kasutajad ei oma sellise kontrolli teostamiseks vajalikke pädevusi, saavad nad selles osas toetuda usaldusväärsetele ekspertidele. Avatud lähtekoodiga süsteemi toimimist näitlikustavad sellised projektid nagu Linux ja Wikipedia.

Neid kolme omadust omavate süsteemide detailsem kirjeldamine ületaks käesoleva artikli piirid. Laas (2023) näitab, et Hassan ja Sallah (2019) pakuvad välja bloki ahelal põhineva autentimissüsteemi, millel on eespool toodud omadused. Postfenomenoloogia seisukohalt on oluline, et disainiotsused arvestaksid, kuidas kujunevad meie arusaamad iseendast ja maailmast suhtes tehnikaga ja selle vahendusel. Seetõttu ei ole me piirdunud ainult süvavõltsingute poolt esile kutsutud võimetusega eristada informatsiooni, vaid arvestanud inimese hermeneutilise suhte ja tehnilise konteksti mitmetahulisusega, milles sisaldub võimekus kohandada tehnikat süvavõltsingute tõttu lootusetuna tunduv olukorras ja arvestada muutuda võivate hoiakutega. Süvavõltsingud puudutavad enam kui pildil kuvatav ja sellele kogemusele vastavad hermeneutilised strateegiad.

Kokkuvõte

Paljuski on süvavõltsingud saamas osaks digisalvestistest, platvormidest ja olemasolevatest TI kui süsteemiga seotud tehnilistest ja sotsiaalsetest normidest, mis kõik mängivad rolli meie igapäevases kogemuses. Süvavõltsingute kontekstis saab täheldada, et me tõlgendame maailma (pildi)tehnikaga kaudu, kuid oluline on ka, kuidas tõlgendatakse pilditehnikat ja mis muutusi see kaasa toob. Süsteemne vaade koos postfenomenoloogiaga toovad esile, kuidas vaataja kogemuses hakkavad süvavõltsingute tõttu digitaalsed salvestised üha enam sarnanema käsitsi tehtud piltidega.

Nende muutuse tulemusel muutuvad salvestistega seotud episteemilised normid sarnasemaks käsitsi tehtud piltidega seotud normidega, mille kohaselt kujutise tõepärasuse määrab selle päritolu ning allika usaldusväärsus. See muutus annab aimu, mida tuleks teha süvavõltsingutega võitlemiseks. Pakume välja, et lahendus peaks olema pigem sotsiotehniline, mitte pelgalt hoiakuline. Esitame kolm disainiomadust, mis salvestiste tõepärasust kontrollida võimaldaval autentimissüsteemil

¹⁴ Detsentraliseeritud süsteemidel on omad haavatavused. Laas (2023) näitab, et need ei ole selliste süsteemide apriõorseks välistamiseks piisavalt kaalukad.

peaksid olema. Tavakasutaja jaoks toimiksid need heuristikatena, mis aitavad erisuguste tehnoloogiliste lahenduste vahel valida.

Allikad

Achterhuis, Hans. 2001. *American Philosophy of Technology: The Empirical Turn*. Bloomington: Indiana University Press.

Ayyub, Rana. 2018. „I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me: Rana Ayyub.“ – *HuffPost*, 3. november. https://www.huffpost.com/archive/in/entry/deepfake-porn_a_23595592.

Baig, Rachel. 2022. „The deepfakes in the disinformation war.“ – *Deutsche Welle*, märts. <https://www.dw.com/en/fact-check-the-deepfakes-in-the-disinformation-war-between-russia-and-ukrainea-61166433>.

Carlson, Matthew. 2021. „Skepticism and the Digital Information Environment.“ – *SATS* 22 (2): 149–167. <https://doi.org/10.1515/sats-2021-0008>.

Cavedon-Taylor, Dan. 2013. „Photographically Based Knowledge.“ – *Episteme* 10 (3): 283–297. <https://doi.org/10.1017/epi.2013.21>.

Christopher, Niles. 2020. „We’ve Just Seen the First Use of Deepfakes in an Indian Election Campaign.“ – *Vice*, veebruar. https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp.

Coeckelbergh, Mark. 2020. „Technoperformances: Using Metaphors from the Performance Arts for a Postphenomenology and Posthermeneutics of Technology Use.“ – *AI & SOCIETY* 35 (3): 557–568. <https://doi.org/10.1007/s00146-019-00926-7>.

———. 2021. „Time Machines: Artificial Intelligence, Process, and Narrative.“ – *Philosophy & Technology* 34: 1623–1628. <https://doi.org/10.1007/s13347-021-00479-y>.

———. 2022. „The Grammars of AI: Towards a Structuralist and Transcendental Hermeneutics of Digital Technologies.“ – *Technology and Language* 2 (7): 148–161. <https://doi.org/10.48417/TECHNOLANG.2022.02.09>.

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.

Fallis, Don. 2020. „The Epistemic Threat of Deepfakes.“ – *Philosophy & Technology* 34 (4): 623–643. <https://doi.org/https://doi.org/10.1007/s13347-020-00419-2>.

Ford, Bryan. 2021. „Technologizing Democracy or Democratizing Technology? A Layered-Architecture Perspective on Potentials and Challenges.“ – *Digital Technology and Democratic Theory*, toimetanud Lucy Bernholz, Hélène Landemore ja Rob Reich, 274–308. Chicago & London: The University of Chicago Press.

Fried, Samantha J. ja Robert Rosenberger. 2021. *Postphenomenology and imagining: how to read technology*. Lanham: Lexington Books.

Galindo, Gabriela. 2020. „XR Belgium posts deepfake of Belgian premiere linking Covid-19 with climate crisis.“ – *The Brussels Times*, aprill. <https://www.brusselstimes.com/news/belgium-all-news/politics/106320/xr-belgium-posts-deepfake-of-belgian-premiere-linking-covid-19-with-climate-crisis/>.

Govier, Trudy. 1997. *Social Trust and Human Communities*. Montreal: McGill-Queen’s University Press.

- Habgood-Coote, Joshua. 2023. „Deepfakes and the Epistemic Apocalypse.“ – *Synthese* 201 (3): 1–23. <https://doi.org/10.1007/s11229-023-04097-3>.
- Hasan, Haya R. ja K. Salah. 2019. „Combating Deepfake Videos Using Blockchain and Smart Contracts.“ – *IEEE Access* 7: 41596–606. <https://doi.org/10.1109/ACCESS.2019.2905689>.
- Heidegger, Martin. 1989. „Küsimus tehnika järele.“ Tõlkinud Ülo Matjus. – *Akadeemia* 6 (1): 1195–1222.
- Hopkins, Richard. 2012. „Factive Pictorial Experience: What’s Special about Photographs?“ – *Noûs* 46 (4): 709–731. <https://doi.org/10.1111/j.1468-0068.2010.00800.x>.
- Ilhde, Don. 1979. *Technics and Praxis*. Dordrecht: Reidel.
- . 1990. *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
- . 1995. *Postphenomenology: Essays in the Postmodern Context*. Evanston, Ill: Northwestern University Press.
- . 2016. *Husserl’s Missing Technologies*. New York: Fordham University Press.
- Laas, Oliver. 2023. „Deepfakes and trust in technology.“ – *Synthese* 202 (5): 1–40. <https://doi.org/10.1007/s11229-023-04363-4>.
- Matthews, Taylor. 2022. „Deepfakes, Intellectual Cynics, and the Cultivation of Digital Sensibility.“ – *Royal Institute of Philosophy Supplement* 92: 67–85. <https://doi.org/10.1017/s1358246122000224>.
- Rini, Regina. 2020. „Deepfakes and the Epistemic Backstop.“ – *Philosophers’ Imprint* 20 (24): 1–20. <https://www.philosophersimprint.org/020024>.
- Rosenberger, Robert. 2008. „Perceiving Other Planets: Bodily Experience, Interpretation, and the Mars Orbiter Camera.“ – *Human Studies* 31 (1): 63–75. <https://doi.org/10.1007/s10746-007-9078-1>.
- Rosenberger, Robert ja Peter-Paul Verbeek. 2015. „A Field Guide to Postphenomenology.“ – *Postphenomenological Investigations: Essays on Human-Technology Relations*, toimetanud Robert Rosenberger ja Peter-Paul Verbeek, 9–41. Lanham: Lexington Books.
- Schick, Nina. 2020. *Deepfakes: The Coming Infocalypse*. New York: Twelve.
- Strupp, Catherine. 2019. „Fraudsters Used AI to Mimic CEO’s Voice in Unusual Cybercrime Case.“ – *The Washington Post*, august. <https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402>.
- van de Poel, Ibo. 2020. „Embedding Values in Artificial Intelligence (AI) Systems.“ – *Minds and Machines* 30 (3): 385–409. <https://doi.org/10.1007/s11023-020-09537-4>.
- van den Berg, Bibi ja Ester Keymolen. 2017. „Regulating Security on the Internet: Control Versus Trust.“ – *International Review of Law, Computers & Technology* 31 (2): 188–205. <https://doi.org/10.1080/13600869>.
- Verbeek, Peter-Paul. 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Tõlkinud Robert P. University Park, Pa: Pennsylvania State Univ. Press.
- . 2008. „Cyborg Intentionality: Rethinking the Phenomenology of HumanTechnology Relations.“ – *Phenomenology and the Cognitive Sciences* 7 (3): 387–395. <https://doi.org/10.1007/s11097-008-9099-x>.
- . 2019. „Postphenomenology and Ethics.“ – *Technology Ethics: A Philosophical Introduction and Readings*, toimetanud Gregory J. Robson ja Jonathan Y. Tsou, 42–51. New York: Routledge.

Walton, Kendall L. 1984. „Transparency of Pictures: On the Nature of Photographic Realism.“ – *Noûs* 18 (1): 67–72. <https://doi.org/10.2307/2215023>.

Wellner, Galit. 2019. „Digital Cultural Sustainability.“ – *Sustainability in the Anthropocene: Philosophical Essays on Renewable Technologies*, toimetanud Róisín Lally, 135–149. Lanham, Maryland: Lexington Books.

---. 2020. „Material Hermeneutic of Digital Technologies in the Age of AI.“ – *AI & SOCIETY* 38: 2159–2166. <https://doi.org/10.1007/s00146-020-00952-w>.

Wiltse, Heather. 2014. „Unpacking Digital Material Mediation.“ – *Techné: Research in Philosophy and Technology* 18 (3): 154–182. <https://doi.org/10.5840/techne201411322>.

Kevin Rändi – doktorant ja nooremteadur Tallinna Ülikooli humanitaarteaduste instituudis. Lisaks töötab ta külalislektorina Tallinna Tehnikakõrgkoolis. Tema peamised huvid hõlmavad postfenomenoloogiat, tehnikafilosoofiat, protsessifilosoofiat ja kultuurifilosoofiat.

E-post: kevin.randi[at]tlu.ee

Oliver Laas – PhD, filosoofia lektor Tallinna Ülikooli humanitaarteaduste instituudis. Tema uurimisteemad hõlmavad mängu-uuringuid, metafüüsikat, loogikat, tehisintellekti filosoofiat, tehnikafilosoofiat, virtuaalsuse filosoofiat ja semiootikat.

E-post: oliver.laas[at]tlu.ee

The Handmade Deepfake World: A Postphenomenological Study of the Changing Human–Imaging Technology Relation in the Context of Deepfakes

Kevin Rändi, Oliver Laas

Keywords: postphenomenology, deepfakes, artificial intelligence, hermeneutic relation, imaging technology

Artificial intelligence (AI) exerts a transformative pressure on our ongoing engagement with other technologies and on our meaning-making practices as they relate to ourselves as well as the world. Intriguing predictions and narratives about possible future outcomes, however, are not limited to AI as a ubiquitous technology that needs careful ethical consideration. Some relevant issues concern our present-day imaging technologies. Photos, videos, and other recordings have traditionally contributed to the creation of a shared understanding regarding the facts. Machine learning algorithms enable the creation of different forms of synthetic media, including deepfakes, which greatly simplify the manipulation of recordings. Deepfakes are notorious because they enable the generation of believable recordings of events that never took place by, for example, seamlessly replacing human faces and voices. After they emerged sometime around 2017, deepfakes have been used for generating involuntary pornographic material and spreading political disinformation. Such uses may be reason enough for taking legislative actions toward combating deepfakes.

Moral issues notwithstanding, an underlying intellectual concern has to do with the future of trust in recordings as a means of establishing the facts. The ubiquity of AI tools that facilitate the manipulation of recordings in ways that make them indistinguishable from real ones could further exacerbate the problem of disinformation. As such, deepfakes are a serious threat that cause epistemic harm and could possibly bring about a form of dystopia called the ‘infocalypse’ (Schick 2020). This possibly leads to a loss of control over information on a large scale.

We consider deepfakes from a phenomenological point of view that puts special emphasis on the role of technologies (or artefacts) as mediators of how the world gains its meaning (Verbeek 2005), known as postphenomenology. Don Ihde (1934–2024), the originator of postphenomenological thought, emphasised the practice-oriented, both culturally and technologically textured lifeworlds in which humans make sense of their surrounding world, and often use artefacts in diverse ways. Thus, postphenomenological research avoids totalising accounts of the future—like the infocalypse—and instead focuses on the complexities found in the practices and varieties of human–technology relations. Regarding deepfakes, this means taking a socio-technically embedded view of embodied human beings and their engagement with imaging technologies in which machine learning plays an active role in shaping our experience of the images.

From a postphenomenological perspective, the relation between humans and technology that characterises our use of imaging technologies is a hermeneutic one. The world mediated through technological imaging could be more ‘transparent’ for experts than non-experts and relies on strategies for interpreting the image (Rosenberger 2008). On the screens of our everyday life, even though many users are familiar with the technical or platform-related context of recordings, it could nevertheless be

argued that deepfakes further dissolve our ability to interpret and change strategies to distinguish a truthful representation from a deceptive one.

However, a purely postphenomenological account would provide an incomplete understanding of deepfakes and the possibilities of meaning-making through imaging technologies. With the increasing involvement of AI and algorithms in imaging technologies, postphenomenology has recognised the need for expanding hermeneutic human-technology relations (Wiltse 2014; Wellner 2020). Furthermore, AI poses problems for artefact-centred approaches in postphenomenology (Coeckelbergh 2022). In light of this, we argue for a socio-technical systems approach to human-technology relations combined with postphenomenology in the context of AI. From this combined theoretical perspective, we argue for a hermeneutical programme that accounts for how humans can change the perspectives from which they view recordings and how this influences their interpretations of said recordings. This is also significant for ethical purport in design against deepfakes.

Contrary to other philosophical explanations, we argue that deepfakes are epistemically harmful because they undermine trust in recording technologies. This changes our hermeneutic relationship with recordings. We analyse this change by comparing how traditional and digital photography mediates our relationship with the world. Relying on Hopkins (2012), we then suggest that deepfake technology, viewed as an AI system (a subtype of socio-technical systems), changes our hermeneutic relationship to recordings so that it begins to resemble our relationship to handmade images. After analysing the latter, we conclude that the reliability of recordings, in the age of deepfakes, depends on their provenance and the reputation of their source, as is the case with handmade images. In other words, deepfakes change the epistemic norms associated with and our hermeneutic relationship to recordings so that they resemble those we have toward handmade images.

Kevin Rändi is a PhD student and junior researcher in the Studies of Cultures programme at Tallinn University. He also works as a guest lecturer at Tallinn University of Applied Sciences. His primary interests encompass postphenomenology, philosophy of technology, process philosophy, and philosophy of culture.

E-mail: kevin.randi[at]tlu.ee

Oliver Laas has a PhD in philosophy from the School of Humanities in Tallinn University. His research interests include game studies, metaphysics, logic, philosophy of artificial intelligence, philosophy of technology, philosophy of virtuality, and semiotics. He works as a lecturer in philosophy at the School of Humanities in Tallinn University.

E-mail: oliver.laas[at]tlu.ee