

# Tehisintellekti salaplaanid: kuidas valeinfo mängib valimistega



**KATERYNA BOTNAR**  
Eesti Rahvusraamatukogu  
meediapädevuse spetsialist

**S**õnal „valimised“ on potentsiaali saada 2024. aastal kõige kasutatavamaks sõnaks, arvestades inimeste hulka, kes sel aastal valima läheb. Kas me oleme valmis valima ja kas oleme piisavalt teadlikud, et oma häält mitte kaduma lasta? Kõik need küsimused on jätkuvalt aktuaalsed, kuid nüüd on lisandunud uued seoses tehisintellektiga.

Lihtsustades võiks öelda, et tehisaru oskab tõlkida inimfantaasiat teksti-, pildi-, audio- ja videokeelde. Sellele ei ole lihtne vastu hakata, eriti kui pole aimu, mida tehistaip just valimiste kontekstis teha suudab. Sestap oleme kogunud mõned fragmendid ulatuslikust mosaiigist teemal „tehisintellekt ja valimised“, et oskaksite olla ettevaatlikud ja ka teadlikud uutest riskidest.

Soojenduseks soovitame lühikest arutelu neuroteadlase Jaan Aruga Kuku raadio päraselõunasaaates, et viia ennast teemaga kurssi ning tuletada meelde, mis

on tehisintellekt ning millised on tema peamised omadused.

Mõned uurijad püüavad juba praegu välja selgitada ohtusid, mis kaasnevad tehisintellekti hüppelise arenguga. Näiteks räägitakse ekstremismiideede levikust ja viisidest, kuidas tehisaru võib sellesse panustada: näiteks genereerides tekste, mis matkivad inimeste kõnekeelt ja mis on palju paremini koostatud kui Google Translate'i abil loodu. Tehistaip ei piirdu sugugi ainult tekstiloomega, sest pärast Hamasi 7. oktoobri rünnakut märgati paljude juudivaenuliku sisuga võltspiltide laiemat levikut.

Aga kuidas selline info ikkagi liigub? Täpsemini öelda seda levitatakse ja muu hulgas erinevate võrgustike kaudu.

ISD Global on teinud uuringu, milles keskenduti Hiina Spamouflage'i võrgustikule (ISD 2023). Uuringu tutvustuses loeme järgmist: „Hiina Kommunistlikku Parteid toetav võrgustik Spamouflage korraldab veebis mõjutuskampaaniaid, sh kasutades tehisintellekti loodud pilte eesmärgiga lõhestada lääne ühiskonda ning propageerida Ameerika-vastaseid narratiive. Kui varem keskendus võrgustiku jagatud sisu peamiselt Venemaa sõjale Ukrainas, siis alates oktoobrist lisandusid Iisraeli-Hamasi

sõja narratiivid ning praegu Bideni kui sõjast sõltuva inimese portreerimine.“

## VALE-SOTSIAALMEEDIAKONTOD JA VÄÄRNARRATIIVID

Mida Spamouflage'i võrgustik teeb?

- ▶ Kasutab läbisegi päris- ja tehisaru loodud pilte ja tekste, et levitada oma propagandat eesmärgiga mõjutada oma sihtrühmi USAs ja maailmas sotsiaalmeedias loodud võltsprofiilide abil.
- ▶ Loob uusi sotsiaalmeediakontosid, mille profiilipildid on tehisintellekti loodud.
- ▶ Genereerib poliitikute ja presidendi-kandidaatide postreid, nt kujutatakse Joe Bidenit sõjast sõltuvana.
- ▶ Jagab väärnarratiive, näiteks et USA on süüdi Venemaa Ukrainas alustatud sõja pikenedes, kuna ta varustab Ukrainat relvadega.

On näha, et viimasel ajal keskendub Spamouflage üha rohkem USA 2024. aasta valimistele: levitatakse erinevaid kuulujutte poliitikutest, Ukraina abistamisest, piirikriisist ja muust.

Kuid Spamouflage pole sugugi ainukene võrgustik, mis on seotud vale- ja desinfo levitamisega. Nikki on võrgustik, mille moodustab vähemalt 64 X-i (endise Twitteri) kontot (Thomas 2023). Teistest eristab neid see, et kontode sisu on loonud tehisintellekt. Sisu on suunatud Aleksei Navalnõi ja Maria Pevtšihhi halvustamisele. Nikki eesmärk on ohustada Navalnõi ja Pevtšihhi aktsepteerimist Ukrainat toetavate ameeriklaste ja lääneriikide ühiskondade silmis.

ISD ei ole suutnud tõestada, kes on Nikki kampaania taga, kuid teatud ajaline muster on siiski jälgitav: postitused ilmuvad tõistel kellaaegadel Moskva aja järgi.

Mõned kontod kustutas X ära, kuid nende asemele ilmusid kiiresti uued.

Peale tekstilooime seisneb ChatGPT roll veel selles, et tehisaru kasutab üsna tihti vanu teemaviiteid (*hashtag*'e) või mõtleb välja uusi, mille järgi otsides saab

tulemuseks kõik sarnase sisuga postitused Nikki kampaaniast. Inglise keele kasutus on neis tihti kummaline ja kohmakas ehk tegemist on masina kirjutatuga.

Uurijad usuvad, et lähitulevik toob aina rohkem desinfo- ja mõjutuskampaaniaid ning tehisaru areng on sealjuures suureks tõukeks. Näiteks võidakse korraldada kampaaniaid, mis levitavad valeinfot konkreetsetest inimestest, ühiskondlikest ja sotsiaalsetest liikumistest eesmärgiga süvendada inimeste ja gruppide vahelist vastandumist sotsiaalmeedias.

Peamiselt noortele suunatud sotsiaalmeediaplatformil TikTok on samuti omajagu Ukraina maine kahjustamisele suunatud vale- ja desinfot. DFRLabi ja BBC Verify uuriv ajakirjanik Roman Osadtšuk on leidnud tõestusi Ukraina endise kaitseministri vastu TikTakis korraldatud infooperatsioonist, milles on oma osa ka tehisintellektil (Osadtšuk 2023). Nimelt jälgis Osadtšuk TikTakis kahtlaseid kontosid, mille kogu sisu oli vaid üks video. Selle video peategelane oli Ukraina endine kaitseminister Oleksi Reznikov, kes elas luksuslikku elu, ostes kalleid autosid ja mõisu.

Üldjoontes vihjasid kõik videod sellele, et tema ning laiemalt Ukraina valitsus teenivad sõja käigus suuri ebaseaduslikke kasumeid. Videotes kasutati tehisaiibu loodud kõnet. TikTok on tuvastanud ja eemaldanud üle 12 800 sellise konto; neid videoid märgati, kuna X-is ilmusid nende taaspostitused.

Visates niisuguseid narratiive rahvusvahelise publiku ette, püüab Venemaa kahjustada Ukraina imagot ning tekitada hõõrumist nende riikide ja valitsuste vahel, kes Ukrainat aktiivselt abistavad.

## Tasub meeles pidada, et Maailma Majandusfoorumi globaalsete riskide 2024. aasta aruande kohaselt on järgmise kahe aasta suurim risk just vale- ja desinformatsioon.

Kas teadsite, et peaaegu kolm miljardit inimest läheb sel aastal valima, sh riikides nagu India, Mehhiko, Suurbritannia ja USA? Ning probleem ei seisne vaid selles, et

valijatega püütakse enne valimisi manipuleerida, vaid des- ja valeinfo laialdane levik võib kahjustada uute valitsuste mainet. Selle tagajärjed võivad ulatuda skeptilisusest ja umbusaldusest valitsuse suhtes kuni vägivaldsete protestideni (WEF 2024).

Kasvav usaldamatus teabe, meedia ja valitsuste vastu süvendab seisukohtade vastandumist, mis omakorda võib vallandada rahutused ja tekitada ka tõsisid konflikte. Ühiskonnad võivad muutuda polariseerunuks mitte ainult inimeste poliitilise kuuluvuse poolest, vaid ka välisteguritest mõjutatud reaalsustaju osas. Selline suundumus võib olla suur väljakutse sotsiaalsele ühtekuuluvusele ja isegi ühiskonnaliikmete vaimsele tervisele. Emotsioonid ja ideoloogiad varjutavad fakte ehk siis kallutatud info mõjutab uudiste ja faktide tajumist.

Kuidas edasi? Kui riigiasutused püüavad tõkestada valeteabe levikut, tekib samas oht kahjustada meediavabadust. Demokraatlikes riikides on meedial õigus kajastada uudiseid vastavalt oma strateegiale ning riik ei tohi meedia tegevusse üleliia sekkuda.

Euroopa Liit ja riikide valitsused püüavad takistada ja kontrollida tehisaru loodud info levikut, toetada selle tuvastamist ja inimeste teadlikkuse suurendamist.

Selleteemaliste seaduste väljatöötamisega on praegu olukord keeruline, sest protsess vajab rohkem aega, nii et üldjoontes ei ole võimalik arenevale tehnoloogiale kuigi kiiresti järele jõuda.

Mõned meetmed siiski jõustuvad, näiteks OpenAI arendajad on välja töötanud viisi, kuidas piirata ChatGPT väärkasutust (Swenson 2024). Mõnda on rakendatud juba varem, mõni aga on uus algatus, nagu näiteks piltide vesimärgiga tähistamine. Samuti on kavas keelata tehnoloogia kasutamine selliste juturobotite loomiseks, mis kehastaksid poliitikuid, valimistel kandideerijaid ja valitsusliikmeid või levitaksid kallutatud või valeinfot valimistel osalemise kohta.

Nüüdseks on erinevad tehisintellekti kasutatavad programmid võimalised

tootma audiovisuaalseid materjale. The Washington Posti artiklis (Verma ja De Vynck 2024) kirjeldatakse juhtumeid üle maailma, milles poliitikuid või teisi kuulsusi kahtlustatakse kas ebaseaduslikus tegevuses või ebamoraalses käitumises, mille tõestamiseks esitatakse audio- ja visuaalseid materjale.

Vastuseks väidavad kahtlustatavad aga, et tõendid on loonud tehisintellekt. Fikseeritud sai ka valimisaasta esimene roboti poolt Joe Bideni nimel peetud kõne, kus esineja püüdis veenda valijaid mitte osalema valimisaktsioonis „Tuesday primary“. Graphika analüütik Libby Lange, kelle arvamust tihtipeale küsitakse, väidab, et „tehisintellekt destabiliseerib tõe kontseptsiooni ennast, sest kui kõige kohta võib öelda, et see on võltsing, siis polegi enam sellist asja nagu põhitõde“.

Paljude tehisintellekti innustatud suundumuste hulgas on mõned otseselt seotud suure valimisaastaga, näiteks USA iduettevõtte Delphi projekt Chat2024 (Velazco 2024), mis võimaldab esitada juturobotile küsimusi poliitika, poliitikute ja kandidaatide kohta. Robotit on õpetatud osalejate kõnede, intervjuude ja transkribeeritud videote toel, kuid kindlasti ei tohi huviline loobuda omaenda analüüsist.

Vähemalt üks poliitik, Shamaine Daniels, on hakanud kasutama tehisintellektil põhinevat robotit, mis suhtleb valijatega varem kirjutatud vestluste põhjal.

## TEHISINTELLEKT SUUDAB LUUA VÕLTSTEKSTE, -PILTE JA -HELI

Kuidas teie reageeriksite, kui kuuleksite tuttavat poliitikut rääkimas veidrat juttu? Mis tundeid see teie tekitaks ja kuidas mõjuks? Uued sünteesitud hääled on, nagu arvata võib, tehisintellekti toodang. Tehistaipu õpetatakse erinevate häälte näitel ning talle antakse ette tekst, mida ta peab rääkima. Joe Bideni ja Donald Trumpi hääled on juba loodud ning kui soovite kasutada kellegi teise häält, mida pole veel loodud, siis loovad osavad tegijad selle teile lühikese ajaga ja võrdlemisi väikese tasu eest.

Microsoft väidab, et programm VALL-E on samuti võimeline looma kindlat häält ning seda kiiresti, valides oma n-ö raamatukogust võimalikult sarnase ehk siis päris häälega lähedase hääle. Õnneks pole VALL-E teenused veel avalikult kättesaadavad ning teine ressurss, Resemble AI, vajab reaalse inimese antud nõusolekut salvestatud hääle kloonimiseks.

Tõepoolest, oleme praeguseks juba mõnes mõttes harjunud teadmise, et tehisintellekt suudab luua võltstekste, -pilte ja -heli, kuid Matthew Quinlan väidab oma artiklis, et tehistaipu võib kasutada ka sihtrühma määratlemiseks, et jõuda potentsiaalsete valijateni (Quinlan 2023).

Asutuse Centre for Public Impact aruanne toob näiteks 2016. aasta valimised ning kuidas Cambridge Analytica kogunud andmed võimaldasid potentsiaalsetel valijatel saada sihitud reklaami, mis põhines igatühe sotsiaalmeedia kasutamise

hindamisel ehk individuaalsel psühholoogial. Võib oletada, et varsti võib leida uusi näiteid tehisintellekti rakendamiseks sama eesmärgi nimel.

Tehisintellekti mitmed väljendused võivad tekitada üsna erinevaid ja kohati ka vastuolulisi tundeid – kellelegi rõõmu ja kellelegi muret. Kuid meie Eesti Rahvusraamatukogus rõhutame nagu alati, et tehistaibu arenguga tuleb kaasa minna, kasvõi tasapisi. Igaüks oma kiirusega. Meie meediapädevuse spetsialistid tegutsevad selle nimel, et teil oleks aktuaalseim ja värskem info alati käe ulatuses.

*Kui vajate täiendavat infot, nõuannet või soovitusi, palun võtke ühendust [desinfo@rara.ee](mailto:desinfo@rara.ee). Rahvusraamatukogust leiate meediapädevuse ja desinfo seotud kirjandust, samuti uuringuid, valdkonnateemalisi koolitusi ja seminare. Rohkem infot <https://www.rara.ee/uuri/desinformatsioon/>*

---

## KASUTATUD ALLIKAD

- ARU, Jaan. Tehisintellekti maailm. – Intervjuu „Kuku pärastlõunas“. – <https://kuku.pleier.ee/podcast/kuku-parastlouna/148724>
- INSTITUTE FOR STRATEGIC DIALOGUE (ISD), 2023. „Spamouflage“, pro-CCP network, targets West with AI images on Israel-Hamas conflict and shifts to US elections. – 19. detsember. – <https://www.isdglobal.org/isd-in-the-news/pro-ccp-network-capitalises-on-the-israel-hamas-conflict-to-target-western-audiences-shifting-toward-us-elections/>
- INSTITUTE FOR STRATEGIC DIALOGUE (ISD), 2024. How extremists may potentially use AI to further reach, recruitment. – 2. jaanuar. – <https://www.isdglobal.org/isd-in-the-news/how-extremists-may-potentially-use-ai-to-further-reach-recruitment/>
- LEE, Lloyd, 2023. Election campaigns are already using AI with some degree of success. Experts are concerned it will usher in a new era of disinformation. – Business Insider, 26. juuni. – <https://www.businessinsider.com/republican-democratic-political-election-campaigns-using-ai-disinformation-artificial-intelligence-2023-6>
- OSADTŠUK, Roman, 2023. Massive Russian influence operation targeted former Ukrainian defense minister on TikTok. – Digital Forensic Research Lab (DFRLab), 14. detsember. – <https://dfrlab.org/2023/12/14/massive-russian-influence-operation-targeted-former-ukrainian-defense-minister-on-tiktok/>
- QUINLAN, Matthew, 2023. AI is finding its voice and that's bad for democracy. – World Economic Forum, 14. november. – <https://www.weforum.org/agenda/2023/11/ai-is-finding-its-voice-and-that-s-bad-for-democracy/>
- SWENSON, Ali, 2024. Here's how ChatGPT maker OpenAI plans to deter election misinformation in 2024. – AP, 17. jaanuar. – <https://apnews.com/article/ai-election-misinformation-voting-chatgpt-altman-openai-0e6b22568e90733ae1f89a0d54d64139>
- THOMAS, Elise, 2023. „Hey, fellow humans!“. What can a ChatGPT campaign targeting pro-Ukraine Americans tell us about the future of generative AI and disinformation? – Institute for Strategic Dialogue. Digital Dispatches: ISD's blog for the latest in data-driven research on hate, disinformation and extremism. 5. detsember. – [https://www.isdglobal.org/digital\\_dispatches/hey-fellow-humans-what-can-a-chatgpt-campaign-targeting-pro-ukraine-americans-tell-us-about-the-future-of-generative-ai-and-disinformation/](https://www.isdglobal.org/digital_dispatches/hey-fellow-humans-what-can-a-chatgpt-campaign-targeting-pro-ukraine-americans-tell-us-about-the-future-of-generative-ai-and-disinformation/)
- VELAZCO, Chris, 2024. The top tech trends in 2024 to watch. – Washington Post, 8. jaanuar – <https://www.washingtonpost.com/technology/2024/01/08/tech-trends-2024/>
- VERMA, Pranshu ja DE VYNCK, Gerrit, 2024. AI is destabilizing 'the concept of truth itself' in 2024 election. – Washington Post, 22. jaanuar. – <https://www.washingtonpost.com/technology/2024/01/22/ai-deepfake-elections-politicians/>
- WORLD ECONOMIC FORUM (WEF), 2024. The Global Risks Report 2024. 19th edition. Insight Report. – Jaanuar. – <https://tinyurl.com/GlobalRisksRaport>