



25.02.2026

Lausunto

Asia: VNS 7/2025 vp Valtioneuvoston tulevaisuusselonteon ensimmäinen osa Strateginen toimintaympäristöanalyysi sekä skenaarioita vuoteen 2045

AIDEMOC on kuusivuotinen, strategisen tutkimuksen neuvoston rahoittama monitieteinen tutkimushanke. Hankkeessa tutkitaan tekoälyn, ja erityisesti generatiivisen tekoälyn vaikutuksia yhteiskunnalliseen kehitykseen erityisesti demokraattisten prosessien näkökulmasta.

Kiitän tulevaisuusvaliokuntaa mahdollisuudesta lausua tekoälyn hyödyntämisen uhkista ja mahdollisuuksista informaatiovaikuttamisen ja yhteiskunnallisen vakauden näkökulmista. Lausun tulevaisuusselonteosta (VNS 7/2025) seuraavaa:

1. Yhteiskunnallinen vakaus

Yhteiskunnallisen vakauden keskeisiä sosio-kognitiivisia mekanismeja ovat yhteenkuuluvaisuus, yhteistyöhalukkuus ja sosiaalinen koheesio. Niiden edellytyksiä ovat kokemus stabiilista ja turvallisesta ympäristöstä (Hersh & Stapleton 2006, Gaffney ym., 2025), viranomaisiin ja muuhun yhteiskuntaan kohdistuvasta luottamuksesta, oikeudenmukaisuudesta, osallisuudesta ja oikeuksien yhdenvertaisesta toteutumisesta (mm. Hersh & Stapleton 2006).

Kasvava tutkimusnäyttö osoittaa, että tekoälyn avulla voidaan tukea yhteiskunnallista vakautta, jos teknologiaa kehitettäessä huomioidaan ns. planetaarisesti kestävä kehityksen periaatteet (Gaffney ym., 2025), viranomaisiin kohdistuvan luottamuksen merkitys sekä hyvän hallinnon ja perusoikeuksien yhdenvertainen toteutuminen, sekä taataan kansalaisille riittävä osallisuus.



25.02.2026

Myönteinen kehitys edellyttää lisäksi, että tekoälyteknologioiden hyödyntämiseen liittyvät riskejä ja uhkia tunnistetaan ja ennaltaehkäistään. Näitä ovat mm. teknologiin riippuvuuksiin liittyvät uhat (kohta 2), kansalaisten ja valtion välisen luottamuksen rapautuminen ja tekoälyn negatiiviset vaikutukset hyvinvointiin (mm. julkisten palvelujen heikentyminen), tekoälyn käytöstä johtuva informaatioympäristön saastuminen (ks. kohta 3), sekä mm. demokraattisten prosessien vaarantuminen algoritmisman manipulaation tai suostuttelun vuoksi (kohta 4). sekä taloudellisen ja sosiaalisen hyvinvoinnin heikkeneminen.

2. Yhteiskunnallinen vakaus, geopoliittinen tilanne ja teknologinen riippuvuus

Geopoliittisesti heikentynyt turvallisuustilanne heijastuu myös yhteiskunnalliseen vakauteen. Se nostaa esiin myös mm. kysymykset teknologisista riippuvuuksista, tekoälyn omistussuhteista ja omistussuhteiden kautta siirtyvästä vallasta (Coeckelberg 2026, Zuboff 2019). Erityisen kriittisiä ovat kysymykset, millaisia riippuvuuksia on mm. viranomais-, julkisen sektorin palvelutuotannon ja valtionhallinnon järjestelmillä. Näillä järjestelmillä toteutetaan kansalaisille lakisääteiset palvelut ja taataan perusoikeuksien toteutuminen. Jos viranomaisjärjestelmät pettävät, pääsy palveluihin vaikeutuu tai estyy, ja ihmisten elämänlaatu heikkenee, se murentaa viranomaisiin kohdistuvaa luottamusta sekä yhteiskunnallista vakautta.

Kyse ei ole vain siitä, käytetäänkö julkisissa palveluissa jatkossa kielimallia X tai pilvipalvelua Z. Kyse on siitä, että Suomen kaltaisissa pitkälle digitalisoituneissa maissa tekoälyratkaisut integroidaan asteittain jo olemassaoleviin palvelu-, tietohallinto- ja toiminnallisiin rakenteisiin. Tekoälykomponenttien integroimiseen liittyy monimutkaisia kysymyksiä siitä, missä määrin uudet komponentit lisäävät tai muuttavat mm. kansalaisten henkilötietoja sisältävien viranomaisten tieto-, tiedonhallinta- ja turvajärjestelmien tai niiden taustalla olevan infrastruktuurin jo olemassaolevia riippuvuuksia, sekä myös kysymyksiä käyttöliittymistä, niiden toiminnallisuuksista ja toiminnallisuuksien keskinäisriippuvuuksista.

Rakenteiden monimutkaisuudelle on monta syytä. Sen taustalla on pidempi kehitys, johon heijastuu osittain toisilleen vastakkaisia intressejä (hallinnon avoimuus, yhteentoimivuus ja mm. turvallisuus). Riippuvuus tietyistä ohjelmistotaloista on samalla tavalla monen tekijän summa. Oletettavasti mm. julkisen sektorin digitalisaation



25.02.2026

ohjausmekanismien yhteisvaikutukset mm.hankintasäätelyn kanssa ovat yksi selittävä tekijä. Suomessa on jo jonkin aikaa ohjattu mm. valtion ja julkisen hallinnon ICT-hankintoja yksityiselle sektorille. Eräs lähestymistavan sivuvaikutuksista on, että sivuvaikutus näyttäisi olevan hankintojen keskittyminen isoille toimijoille tavalla, joihin sisältyy myös riski riippuvuudesta,

Tekoälyteknologioiden kohdalla on jatkossa hyvä tunnistaa riippuvuudesta seuraavat riskit, ja etsiä ratkaisuja, joilla estetään uusien, entistä vakavampien riippuvuuk-sien syntyminen ohjausmekanismien ja – ratkaisujen yllättävinä sivuvaikutuksina. Kokonaistilanteen korjaaminen voi olla kuitenkin haastavaa ja vaatia pitkäjänteistä työtä. Tietohallinnon näkökulmasta suomalaisen julkisen sektorin rakenne on tavat-toman monimutkainen, jolloin äkkinäisiin korjausliikkeisiin sisältyy myös aina riski toimintoja ylläpitävien rakenteiden rikkoutumisesta.

Kyse on myös yhteiskunnallisen vakauden näkökulmasta kriittisten järjestelmien huoltovarmuudesta. Radikaaleilla toimenpiteillä voidaan pahimmillaan vaarantaa vi-ranomaisjärjestelmien toiminta siten, että kansalaisten pääsy lakisääteisiin palvelui-hin vaikeutuu. Kustannusvaikutukset voivat osoittautua myös korkeiksi, koska siir-tymä nykyisistä järjestelmistä uusiin voi vaatia myös henkilöstön täydennyskoulu-tusta, toimintojen uudelleenorganisointia ja joissain tapauksissa mahdollisesti myös lainsäädännön päivittämistä.

3. Informaatiovaikuttaminen ja informaatioympäristön saastuminen

Teknologisen riippuvuuden lisäksi mm. erityisesti generatiiviseen tekoälyyn liittyvä informaatioympäristön saastuminen on tutkimuksen näkökulmasta merkittävä yhteis-kunnallista vakautta heikentävä tekijä (Vaccari & Chadwick 2023). Tutkijoiden mu-kaan esimerkiksi koneellisesti generoitu mis- ja disinformaatio voivat vahingoittaa koko tiedollista ympäristöä estämällä nimenomaan paikkansa pitäviä uskomuksia (Fallis 2021). Ne eivät siis johda vain välitetyn sisällön paikkansapitävyyden kyseen-alaistamiseen, vaan tieto altistumisesta on itsessään psykologisesti vahingollista (Fallis 2021; Ecker ym. 2020). Seurauksena voi olla tiedollisesti myrkyttynyt ilma-piiri, joka rapauttaa yhteistyön, kollektiivisen päätöksenteon ja yhteisöllisyyden kog-nitiivisia ja sosiaalisia mekanismeja. Tämä puolestaan voi vaarantaa laajalti mm.



25.02.2026

yhteiskunnan keskeisiä toimintaedellytyksiä (Vaccari & Chadwick 2020; Balliet & Van Lange 2013).

Informaatioympäristön saastumista on useaa eri lajia, ja sitä tapahtuu eri syistä. Esimerkiksi ns. *tahallisessa saastuttamisessa* on kyse generatiivisella tekoälyllä tehtyjen sisältöjen, kuten syvävääreennösten (engl. deepfakes) pahantahtoista ja tarkoitushakuisesta hyödyntämisestä mm. disinformaatiokampanjoissa ja muussa viestinnän vääristelyssä. *Tahattomalla informaatioympäristön saastumisella* puolestaan viitataan mm. kielimallien taipumukseen hallusinoita esimerkiksi siten, että LLM tuottaa tekstin, joka ei ole uskollinen mallin syöteaineistolle (engl. "faithfulness") tai joka ei vastaa todellisuutta (engl. "factuality") (Huang ym. 2023).

Tutkimuksen mukaan mm. suuret kielimallit voivat tuottaa erittäin uskottavaa väärää sisältöä, erityisesti silloin kun sisältö on personoitua tai räätälöity käyttäjän identiteetin mukaan (Gabriel ym. 2024; Menz ym. 2024). Toisaalta erilaiset tekniset, algoritmiset keinot tunnistaa ja hillitä hallusinaatiota ja disinformaatiota kehittyvät nopeasti (mm. Kuznetsova et al. 2025, Costello et al. 2024; Lu 2025). Menetelmien laatu on kuitenkin toistaiseksi vaihtelevaa: tarkkuus riippuu kehotteiden muotoiluista, aihealueesta ja kielestä, ja erityisesti kielimallien kohdalla suorituskyky heikkenee kulttuurisesti kiistanalaisissa tai ei-englanninkielisissä yhteyksissä (Deiana ym. 2023; Kumar ym. 2024). Vaikka tekniset torjuntakeinot ovat lupaavia, ne ovat yhä hauraita. Personoidut korjaukset voivat parantaa erottelukykä, mutta mallien turvamekanismit eivät ole pitkäikäisiä eivätkä tasaisen tehokkaita (Menz ym. 2024).

Informaatioympäristöä voidaan saastuttaa myös *tukahduttamalla* sitä AI-liejulla (engl. AI Slop). AI-liejuksi kutsutaan tyypillisesti generatiivisella tekoälyllä usein mekaanisesti tuotettua ja levitettyä "turhaa" sisältöä. Liejun määrän yleistessä rajusti se täyttää informaatioympäristön, ja hukuttaa alleen merkityksellisen informaation. Tietyissä erityistapauksissa, kuten ns. tiedeliejun kohdalla, se voi tukkia tiedonmuodostuksen kanavat tavalla, jolla on vakavia yhteiskunnallisia seurauksia. Esimerkiksi monet tiedejournaalit ja konferenssit ovat raportoineet, että arvioitaviksi lähetettävien käsikirjoitusten määrä on moninkertaistunut viimeisten kahden vuoden aikana osittain tiedeliejun räjähdysmäisen kasvun myötä. Vaikka kyse ei aina olisikaan varsinaisesta tiedevilpistä, pelkkä koneellisesti generoitujen ja sisällöllisesti merkityksettömien käsikirjoitusten määrän raju kasvu hidastaa tieteelle tärkeää



25.02.2026

vertaisarviointia, sekä saattaa heikentää myös sen laatua. Seurauksena voi olla mm. tieteellisen tiedonmuodostuksen hidastuminen, laadun heikkeneminen ja tutkimustulosten hyödyntämisen hidastuminen.

4.Tekoälyavusteinen taivuttelu: vaalimainonta

Informaatioympäristön saastumisen lisäksi tutkijat ovat aivan hiljattain kiinnittäneet huomiota erityisesti tiettyjen tekoälysovellusten, kuten kielimallipohjaisten chatbottien tehokkuuteen ihmisten asenteiden ja valintojen ohjaamisessa. Yhteiskunnallisen vakauden näkökulmasta erityisen kiinnostavia ovat tutkimustulokset erilaisten tekoälysovellusten manipulatiivisista tai taivuttelevista piirteistä.

Esimerkiksi joulukuussa 2025 Naturessa ja Sciencessä julkaistiin tutkimustuloksia, joissa käsiteltiin LLM-pohjaisten chatbottien kykyä taivutella äänestäjiä (Lin ym, 2025, Hackenburg 2025). Tulosten mukaan lyhyetkin, 5-6 minuutin pituiset keskustelut LLM-pohjaisten chatbottien kanssa voivat taivutella äänestäjän vaihtamaan ehdokasta tehokkaammin kuin perinteiset vaalimainonnan keinot. Tutkijat havaitsivat myös, että erityisesti asiapitoiset, kohteliaat ja miellyttävät botit olivat tehokkaampia taivuttelijoita kuin henkilöihin, tunteisiin tai asiattomuuksiin keskittyvät. Kääntöpuolelana kuitenkin on, että mitä miellyttävämmäksi ja avuliaammaksi botti esiohjelmoidaan, sitä epätarkempia ja hallusinaatioherkempiä ne yleensä ovat (Hackenburg 2025, Chen ym., 2025).

Yhteiskunnallisen vakauden näkökulmasta tulokset kertovat ennenkaikkea siitä, että tekoälyteknologioiden vaikutuksissa kyse ei ole vain teknologiasta, vaan ihmiskäytäjän ja teknologian välisestä vuorovaikutuksesta. Ihminen ei ole vain passiivinen käytön kohde, vaan osapuoli, joka reagoi ja tulkitsee tiedollisesti koneen toimintaa tavoilla, joka mahdollistaa käyttäytymisen ohjaamisen. Kyse ei ole marginaalisesta ilmiöstä, vaan tekijöistä, jotka koskettavat suurinta osaa tekoälyteknologioiden käyttäjistä. Kasvava tutkimusnäyttö nimittäin osoittaa, että suurin osa käyttäjistä yliarvioi tekoälyjärjestelmien suorituskykyyn ja virheettömyyteen, sekä yliarvioi oman kykynsä tunnistaa koneiden tekemiä virheitä ([Shekar et al., 2024](#), Fernandez ym. 2025). Tämä tekee meistä haavoittuvia tekoälyavusteiselle manipuloinnille ja suostuttelulle tavoilla, joihin vasta nyt olemme havahtumassa.



25.02.2026

Anna-Mari Wallenberg

Konsortiojohtaja, STN AIDEMOC (2024-2030)

Kognitiotieteen dosentti, laskennallisten ihmistieteiden yliopistonlehtori

Kognitiotiede, digitaalisten ihmistieteiden osasto

Helsingin yliopisto

sähköposti: anna-mari.wallenberg@helsinki.fi

puhelin: 02 941 29675