

strateginen TUTKIMUS



**AI demokratiaa tukemassa
(Aldemoc)**

Tilannekuvaraportti 2025



SUOMEN AKATEMIA

1. Tiivistelmä

Aldemoc tutkii, miten tekoäly vaikuttaa demokratian tulevaisuuteen, miten digitaalinen demokratia voi kehittyä lyhyellä ja pitkällä aikavälillä, sekä millaista tietoa tarvitaan poliittisiin päätösten ja hyvän hallinnon tukemiseksi. Hankkeessamme selvitämme, miten demokratiaa voidaan tukea tekoälyteknologioiden avulla, sekä kuinka voidaan ennaltaehkäistä tekoälyn käytöstä johtuvaa demokraattisiin instituutioihin ja niiden rakenteisiin kohdistuvan luottamuksen murenemistä. Tutkimme, kuinka tekoälyn käytöllä voidaan rapauttaa luottamusta tahallisesti mm. saastuttamalla informaatioympäristöjä tahallisesti (esim. viestinnän kaappaukset ja aseistaminen) sekä tahattomasti (esim. kielimallien hallusinointi). Lisäksi kehitämme osallisuutta mahdollistavia teknologisia ratkaisuja, kuten mm. ikäänntyvien ja vieraskielisten asioimista tukevia ääni- ja kieliteknologisia sovelluksia. Selvitämme myös, kuinka kansalaisvaikuttamista voidaan tukea demtech- alustoilla.

2. Yhteiskunnallinen haaste

AIDEMOC tutkii tekoälypohjaisten viestintäteknologioiden nopean leviämisen vaikutuksia demokratiaan, erityisesti Suomen ja muiden eurooppalaisten yhteiskuntien kontekstissa. Tuoreiden raporttien mukaan (OECD 2024) luottamus demokraattisiin käytäntöihin on näissä maissa laskussa. Eräs syy on, että julkisten palvelujen heikkeneminen haurastuttaa yhteiskunnallista vakautta (Goldstone 2019), eivätkä ihmiset myöskään koe olevansa osallisia päätöksenteossa.

Monet toivovat, että nykyaikaisen tekoälyn, ja erityisesti generatiivinen tekoälyn, avulla voitaisiin tarjota *kustannustehokkaita saavutettavia palveluita* nopeasti kasvaville erityisryhmille (esim. ikäihmiset, lukutaidottomat maahanmuuttajat, lukihäiriöiset). Uusien tekoälyavusteisten alustojen avulla pyritään myös vahvistamaan *demokraattisia kansalaisosallistumisen ja kollektiivisen päätöksenteon käytäntöjä* (Weyl ym., 2024; Fuller ym., 2024).

Algoritmit herättävät kuitenkin myös huolta (Buhmann & Fieseler, 2023). Ne eivät aina edistä tasa-arvoa, yhteiskunnallista kehitystä tai muita demokraattisia hyveitä (Coeckelbergh 2024, Nemitz 2018). Useiden arvioiden mukaan generatiiviseen tekoälyyn liittyvä mahdollinen informaatioympäristön saastuminen on yksi vakavimmista uhista demokratialle (OECD 2024; EU 2024; UNESCO 2023). Se voi horjuttaa demokraattisia perusmekanismeja heikentämällä kollektiivista päätöksentekoa, yhteistyöhalukkuutta ja luottamusta valtion ja kansalaisten välillä (Vaccari & Chadwick, 2021; Ecker ym., 2021; OECD, 2024). Lisäksi se voi lisätä polarisaatiota, uhata yksilön autonomiaa ja aiheuttaa turvallisuusriskejä (Weidinger ym., 2022; MIT AI Risk Inventory 2024).

2.1. Aldemoc, DEMOC ohjelma ja vaikuttamisen tavoitteet:

Yhdessä muiden DEMOC-ohjelman hankkeiden kanssa Aldemocin tavoitteita ovat:

2.1.1. Sosioteknisten skenaarioiden laatiminen (ennakoinnin tuki):

Vaikuttavan politiikan ja päätöksenteon näkökulmasta keskeisin haaste on tulevien riskien ja mahdollisuuksien tunnistaminen sekä teknologian että yhteiskunnan näkökulmista. DEMOC-ohjelman muiden hankkeiden (mm. STRONG) kanssa Aldemocin tavoitteena on tunnistaa teknologisia ja yhteiskunnallisia mm. demokratian kehitykseen, yhteiskunnalliseen vakauteen ja resilienssiin, globaalien uhkien, kriisien ja digitalisaation vaikutuksiin liittyviä trendejä ja tekijöitä lyhyellä (2030) ja pitkällä (2050) aikavälillä. Skenaarioiden rakentamisessa hyödynnetään mm. laskennallisia malleja, sekä erilaisia aineistonkeruumenetelmiä (mm. Threatcast-metodi). Skenaarioiden avulla jäsennetään trendejä, sekä hahmotetaan tarpeita varautua mahdollisiin tekoälyn demokratialle aiheuttamiin uhkiin. Lisäksi niiden avulla pyritään suuntaamaan tulevaisuutta tekoälyn tarjoamien mahdollisuuksien hyödyntämiseksi.

2.1.2. Tietoekosysteemin saastumisen ehkäisy (disinformaatio ja hallusinaatiot):

Tekoälypohjainen informaatioympäristön saastuminen on yksi vakavimmista uhkista demokratialle (Caldwell ym., 2020; OECD 2024). Aldemocin tavoitteena on selvittää keinoja kahden erityyppisen informaatioympäristön saastumisen vaikutusten ehkäisemiseksi ja hillitsemiseksi: tahallinen (esim. kommunikaatiokaappaukset, algoritmien aseistaminen) ja tahaton mis- ja disinformaation leviäminen (esim. hallusinoivien generatiivisten mallien käyttö).

2.1.3. Digitaalinen demokratia ja julkiset palvelut:

Aldemocissa selvitetään muiden DEMOC-ohjelman hankkeiden kanssa, kuinka tekoälyn avulla voitaisiin parantaa ja tehostaa julkisia palveluja ja hallinnollisia prosesseja, osallisuutta sekä ehkäistä tekoälyn käyttöönottoon liittyviä riskejä (Krepps ym., 2023).

2.1.4. Demokratian kehityksen tukeminen (digitaalinen deliberatiivinen demokratia)

Tuoreet raportit (OECD 2024) osoittavat, että monissa yhteiskunnissa, kuten Suomessa, ihmiset kokevat vaikutusmahdollisuutensa päätöksentekoon heikoiksi. Tutkimukset viittaavat siihen, että tekoälypohjaiset digitaalisen demokratian teknologiat (Tsai ym., 2024; García ym., 2023; Fuller ym., 2024) voivat vahvistaa kansalaisten osallistumista (Tseng 2022). Aldemocin tavoitteena on yhdessä mm. OBaMa ja STRONG-hankkeiden kanssa selvittää, millaisten institutionaalisten käytäntöjen avulla algoritmeja

voitaisiin käyttää kansalaisten osallistamisen sekä demokraattisen päätöksenteon tukena (Poblet 2019; OECD 2021; Jämsen ym., 2022).

3. Tutkimuksen nykytila (state of the art)

3.1.1. Hankkeen asemointi kansainvälisessä tutkimuskentässä

Aiempi tutkimus osoittaa, että demokratia edellyttää kansalaisosallistumista, luottamusta poliittisiin instituutioihin ja halukkuutta yhteistyöhön (Goodin 2008; Goldstone ym., 2019). Monitieteinen näyttö korostaa myös, että luotettavat ja oikeudenmukaiset hallintokäytännöt (Goldstone ym., 2019; Krepps ym., 2023), toimivat ja tehokkaat julkiset palvelut (Perry 2007; Goldstone ym., 2019) sekä pääsy yhteiskunnalliseen kehitykseen tukevat demokratian keskeisiä elementtejä, kuten tasa-arvoa, yhteiskunnallista koheesiota ja vakautta (Perry 2007).

Tekoäly tuo mukanaan sekä riskejä että mahdollisuuksia demokraattisille käytännöille ja hallinnolle (Buhmann & Fieseler, 2023; Coeckelbergh 2024). Esimerkiksi generatiivinen tekoäly voi vahvistaa kansalaisten osallistumista, sitoutumista ja saavutettavia julkisia palveluita (Krepps ym., 2023), mutta se herättää myös huolta syrjinnästä, vastuullisuudesta ja tiedon saastumisesta (Weidinger ym., 2022; MIT AI Risk Repository 2024).

Aiemmassa tutkimuksessa syrjinnän ja vastuullisuuden tapaiset teemat ovat korostuneet. Ihmisryhmien alttiutta uudenlaisiin tekoälypohjaisiin vihamielisiin dis- ja misinformaation muotoihin (Johansson ym. 2023) sekä yksilöllisiä eroja reagoinnissa ei ole juurikaan selvitetty. Olemassaoleva tutkimus tekoälypohjaisesta misinformaatiosta viittaa, että kognitiiviset ja emotionaaliset tekijät selittävät alttiutta misinformaatioon paremmin kuin disinformaatiokampanjoiden vaikuttavuus (Pamment ym., 2019; Ariely 2023). Kognitiivisia mekanismeja, jotka selittävät yksilöllisiä eroja näissä tekijöissä – kuten epistemistä valppautta – ymmärretään kuitenkin vielä huonosti (Goupil ym., 2021).

AIDEMOCin monitieteinen, ja osittain uudenlainen menetelmällinen lähestymistapa sijoittaa sen tutkimuksen eturintamaan useilla alueilla. AIDEMOC laajentaa olemassa olevaa kansainvälistä tutkimusta alueille, kuten teknologiavälitteisen kommunikaatiokaappauksen, hallusinoinnin ja muiden digitaalisten uhkien muotoihin, joita ei ole juurikaan tutkittu yhdistämällä ihmis-, kognitio- ja viestinnän tutkimusta. Lisäksi hanke avaa tuoreen näkökulman puhesynteessin ja ihmisen ja koneen vuorovaikutuksen (HCI) tutkimukselle. AIDEMOC pyrkii nostamaan Suomen eettisesti suunniteltujen ei-ihmismäisten synteettisten äänien kehityksen kärkeen – erityisesti niiden yhdistämisessä hallusinaatioalttiisiin, suuriin kielimalleihin ja niiden hyödyntämiseen. AIDEMOC avaa myös monitieteisen tutkimuskentän, jossa

tarkastellaan empiirisesti digitaalisten demokratia-alustojen kognitiivista käytettävyyttä. Samalla se laajentaa olemassa olevaa kansainvälistä tutkimusta digitaalisten deliberatiivisten teknologioiden ja tekoälyn etiikan saralla uusille alueille, kuten käyttäjälähtöiseen teknologia- ja informaatiomuotoiluun.

4. Monitieteinen yhteistyö

Tekoälyn vaikutus demokratian tulevaisuuteen, digitaalisiin uhkiin ja tekoälyn aseistamiseen sekä tekoälypohjaisten ratkaisujen sosio-kognitiiviset piirteet ovat monimutkaisia ilmiöitä, joissa yhdistyvät teknologiset, yhteiskunnalliset, eettiset ja kognitiiviset näkökulmat. Ne ovat läpileikkaavia teemoja kaikissa DEMOC-ohjelman hankkeissa, sekä myös yhteisiä muiden STN-ohjelmien hankkeiden (mm. DataLit; Repair; SHIELD-ohjelma) kanssa.

AIDEMOC yhdistää hallinto- ja yhteiskuntatieteiden, kognitiivisten ja tietojenkäsittelytieteiden sekä digitaalisen demokratian asiantuntemuksen tutkiakseen demokratian tulevaisuutta algoritmien aikakaudella. AIDEMOC yhdistää myös viestinnän hallinnan, median, riskien ja disinformaation tutkimuksen asiantuntemusta käsitelläkseen monimuotoisia digitaalisia uhkia. Hanke hyödyntää empiiristen ja kokeellisten tieteiden asiantuntijoita yhdistämällä heidän tilastollisen analyysin ja mallintamisen osaamisensa disinformaatiotutkimukseen kerätäkseen, analysoidakseen ja mallintaakseen empiiristä aineistoa piilevistä alttiuksista misinformaatiolle.

5. Vuorovaikutus ja vaikuttavuuden edistäminen

Aldemoc-hankkeen vuorovaikutuksen tavoitteena on tukea demokratian kehitystä tekoälypohjaisten viestintäteknologioiden nopean yleistymisen keskellä. Hanke toteuttaa vaikuttavaa vuorovaikutusta useiden sidosryhmien (mm. ministeriöt, virastot, kansalaisjärjestöt) kanssa eri tasoilla. Vuorovaikutuksen tavoitteena parantaa kansalaisten osallistumismahdollisuuksia, kehittää saavutettavia julkisia palveluja sekä ehkäistä informaatioympäristön saastumista.

Aldemocin on mahdollista tehdä yhteistyötä DEMOC-ohjelman muiden hankkeiden (kuten STRONG ja OBaMa) sekä sidosryhmä- ja vuorovaikutuspartnerien kanssa mm. seuraavissa teemoissa:

Sosioteknisten skenaarioiden laatiminen (ennakointi): Tavoitteena on tuottaa laajassa sidosryhmäyhteistyössä (mm. ministeriöt, YLE, Sitra, virastot) tietopohjaa tulevaisuuden politiikkaa ja päätöksentekoa varten tunnistamalla tekoälyn ja yhteiskunnallisen kehityksen riskejä ja mahdollisuuksia vuosille 2030 ja 2050.

Informaatioekosysteemin saastumisen ehkäisy: Hankkeessa selvitetään keinoja hallusinaatioiden ja disinformaation torjumiseksi sekä tahallisen että tahattoman väärän tiedon leviämisen osalta yhteistyössä mm. relevanttien viranomaisten (mm. Traficom, VM, PLM, Tukes, Vero) sekä media- ja viestintätoimijoiden kanssa (mm. YLE).

Digitaalinen demokratia ja julkiset palvelut: Osassa hankkeen työpaketteja kehitetään vuorovaikutuskumppanien kanssa julkisten palveluiden prototyyppisovelluksia, joilla kerätään järjestelmällistä tutkimustietoa erilaisten tekoälytekniikoiden soveltuvuudesta, hyödyllisyydestä, riskeistä ja lainsäädännöllisistä kysymyksistä. Lisäksi hankkeessa on tarkoitus tuottaa tietopohjaa julkisen sektorin saavutettavuuden ja tehokkuuden parantamiseen erityisesti heikommassa asemassa oleville ryhmille yhteistyössä viranomaisten (mm. DVV) kanssa.

Algoritminen osallistaminen: Tavoitteena on tuottaa suosituksia siitä, kuinka tekoälyä voidaan hyödyntää kansalaisten osallistamisen ja demokraattisen päätöksenteon tukena.

- **Sidosryhmäyhteistyö ja viestintäkanavat**

Aldemoc tekee aktiivista sidosryhmäyhteistyötä päätöksentekijöiden, viranomaisten, kansalaisjärjestöjen, median sekä tutkijoiden kanssa. Hankkeen vuorovaikutuksesta vastaa luova toimisto Sherpa, joka koordinoi myös viestintää.

Keskeisiä toimia:

- **Tapahtumat ja työpajat:** Hankkeen yhteydessä järjestettävissä tapahtumissa kerätään rekisteriä kiinnostuneista toimijoista. Heille suunnataan uutiskirje, jossa tiedotetaan tutkimuksen edistymisestä, ajankohtaisista aiheista ja tulevista tapahtumista.
- **Sosiaalinen media ja videot:** Hankkeella on aktiivinen LinkedIn-kanava (<https://www.linkedin.com/company/aidemoc/>) ja YouTube-tili (<https://www.youtube.com/@Aldemoc-hanke>), jonne tuotetaan asiantuntijahaastatteluja, lyhytvideoita ja tapahtumataltiointeja. Esimerkkeinä mainittakoon esittelyvideo ja kickoff-tapahtuman pikahaastattelut.
- **Kohdennettu mediayhteistyö:** Hankkeen näkyvyyttä pyritään maksimoimaan laadukkaissa medioissa (kuten YLE ja HS) suunnitelmallisella mediapitchauksella, joka huomioi yhteiskunnallisesti ajankohtaiset aiheet.
- **Dialogisarja 2025–2026:** Yhteistyössä Valtioneuvoston kanslian kanssa on suunniteltu dialogisarjaa tarjoamaan areena syvälliselle keskustelulle tekoälyn ja demokratian suhteesta.

Kaiken viestinnän yhteydessä pyritään tuomaan esiin tutkimustuloksia ymmärrettävästi ja vaikuttavasti, erityisesti päätöksenteon ja

kansalaiskeskustelun tueksi. Tiedotuksessa mainitaan mahdollisuuksien mukaan Suomen Akatemia ja STN. Hankkeen virallinen aihetunniste on **#Aldemoc**.