

Kirsi Hantula

TIEDOLLISET OIKEUDET TEKOÄLYN AIKAKAUDELLA

Demokraattisten yhteiskuntien vastalääke
tekoälykehityksen tiedollisen vallan
keskittymiselle

Sitran muistio

© Sitra 2025

**Tiedolliset oikeudet tekoälyn aikakaudella – Demokraattisten yhteiskuntien
vastalääke tekoälykehityksen tiedollisen vallan keskittymiselle**

Kirjoittajat: Kirsi Hantula sekä työryhmä, johon kuuluivat Jukka Vahti, Tiina Härkönen,
Heikki Lauha, Rosa-Maria Mäkelä, Tarmo Toikkanen, Hannu-Pekka Ikäheimo ja
Antti Lehtinen

Taitto: Grano Oy

ISBN 978-952-347-419-2 (PDF) www.sitra.fi

ISSN 2737-1034 www.sitra.fi

Sitran muistiot ovat tulevaisuustyömme taustaksi tuotettuja sisältöjä.

Sisällys

Esipuhe	4
Tiivistelmä	6
Sammanfattning	8
Summary	10
1. Tiedollinen epätasa-arvo on polttava demokratiakysymys	
digitalisoituvissa yhteiskunnissa	12
2. Tiedollinen valta generatiivisen tekoälyn kehittämisessä	
keskittyy teknologiajäteille	16
3. Tekoälykehityksen tiedollisen epätasa-arvon yhteiskunnalliset	
seuraukset	20
4. Tiedolliset oikeudet demokraattisemman tekoälykehityksen	
perustana	24
Pääsy tietoon	24
Osaaminen ja valmiudet	26
Osallistuminen	29
Ei pakkoa tai painostusta	30
5. Jälkisanat: Kolme askelta tiedollisten oikeuksien soveltamiseen	
käytännössä	32
Lähteet	34

Esipuhe

Sosiaalisen median nousu rinnastettiin aikanaan kirjapainotaidon ja sanomalehdistön syntyyn. Nyt samaa tarinaa toistetaan hurjaa tahtia kehittyvistä tekoälysovelluksista.

Vanhan viisauden mukaan teknologisten murrosten vaikutuksia liioitellaan lyhyellä aikavälillä, mutta aliarvioidaan pidemmällä jaksolla. Sosiaalisen median osalta elämme jo keskipitkää aikaväliä, ja joitain yhteenvetoja voidaan jo tehdä. Aiempien viestintäteknologisten murrosten tavoin sosiaalinen media on tuonut useammat ihmiset tiedonvälityksen ja tiedontuotannon piiriin. Samalla se on tuonut kimpun uusia yhteiskunnallisia ongelmia ratkaistavaksi.

Vähäisin noista ongelmista ei ole vallan kasautuminen kouralliselle kansainvälisiä jättiyrityksiä. Kun somehype oli suurimmillaan, sadat miljoonat käyttäjät ympäri maailman – minäkin – luovuttivat huolettomasti, kuin unissa kävellessä, henkilökohtaisia tietojaan surutta Facebookin ja muiden someyritysten käyttöön. Se synnytti kokonaan uuden ansaintalogiikan, datatalouden, jossa yksilöiden ja yhteisöjen henkilötiedoista ja käyttäytymisen seuraamisesta tuli kauppatavaraa. Nyt monet näistä samoista yrityksistä ovat keskeisiä pelureita tekoälymurroksessa, koska niillä on käytettävissään sekä hurjat määrät tekoälymallien kouluttamiseen tarvittavaa dataa että kasoittain rahaa mallien kehittämiseen.

Tekoälymurros etenee tilanteessa, jossa sosiaalisen median nousun esiin tuomat ongelmat ovat vielä pitkälti ratkaisematta. Virheitä ei kuitenkaan tarvitse toistaa. Jos sosiaalisen median ja datatalouden nousu on jotakin opettanut, niin ainakin sen, että yhteiskunnan läpäisevien teknologioiden tulisi olla nykyistä vahvemmin demokraattisen päätöksenteon piirissä. Tekoälyn osalta se tarkoittaa muun muassa sitä, että hyväksyttävän tekoälyn käytön rajat esimerkiksi julkisissa palveluissa täytyy altistaa julkiselle keskustelulle ja puntaroinnille.

Tämä muistio on jatkoa Sitran aiemmalle työlle digitaalisesta vallasta ja digitalisaation vaikutuksista yhteiskunnallisiin valtasuhteisiin. Tarkoituksena on pohtia, miten yhteiskunnan eri toimijat voisivat hyödyntää tiedollisten oikeuksien tarjoamia näkökulmia tekoälyistyvässä Suomessa



demokraattisemman, yksilön oikeuksia kunnioittavan ja innovatiivisen tietoyhteiskunnan rakentamiseen sekä oman toimintansa kehittämiseen. Samalla muistio taustoittaa Sitran syksyllä 2025 alkavaa kokeilua, jossa toteutetaan yhdessä kumppanien kanssa maailman ensimmäinen kansalaisten luoma tekoälyn säännöstö.

Haluamme myös muistuttaa yhteiskunnan eri sektorien toimijoita siitä, että digitalisaatiokehitys, ja tekoälyn kehittyminen sen osana, eivät ole ”luonnonlakeja”. Meillä on lukuisia keinoja vaikuttaa kehityksen kulkuun – jos haluamme.

Hannu Ikäheimo

Demokratiainnovaatiot-ohjelman johtaja
Sitra

Tiivistelmä

Viime vuosien nopea tekoälykehitys on herättänyt perusteltua huolta, että valta kansainvälisessä tekoälykehityksessä uhkaa keskittyä pienelle joukolle teknologiayrityksiä, joilla on suurimmat teknologiset ja taloudelliset resurssit. Pelkona on, että nämä yritykset pystyvät määräämään tekoälyn kehityksen suunnan, muilta liiemmälti kyselemättä.

Tässä muistiossa tekoälykehitystä käsitellään tiedollisen vallan näkökulmasta. Tiedollisen vallan keskittyminen tekoälykehityksessä on demokraattisille yhteiskunnille yhtä vakava ongelma kuin taloudellisen ja teknologisen vallan keskittyminenkin. Tämä johtuu siitä, että teknologiajätit tekevät tekoälyä kehittäessään monia päätöksiä, joilla on kauaskantoisia ekologisia, eettisiä, sosiaalisia tai yhteiskunnallisia vaikutuksia.

Jotta näitä päätöksiä tehtäisiin jatkossa läpinäkyvämmin ja yhteiskunnan kokonaisuutena puntaroiden, demokraattisten yhteiskuntien tulisi purkaa tekoälyn kehittämiseen pesiytynyttä tiedollista epätasa-arvoa ja vahvistaa kansalaisten tekoälytoimijuutta. Tässä muistiossa esitellään tiedollisten oikeuksien kehikko, joka tarjoaa demokraattisille yhteiskunnille ja erilaisille organisaatioille keinoja päästä työssä alkuun. Alkuperäisen kehikon on laatinut viestinnän emeritusprofessori Hannu Nieminen, mutta kehikkoa on tässä muistiossa muokattu tekoälykehityksen kontekstiin sopivammaksi.

Tiedollisten oikeuksien kehikko auttaa hahmottamaan konkreettisesti, millä tavoin suurille teknologiayrityksille kasautunut tiedollinen valta heikentää yhteiskunnan muiden toimijoiden tekoäly-ymmärrystä ja kaventaa kansalaisten mahdollisuuksia vaikuttaa tekoälyn kehittämisen yhteydessä tehtäviin päätöksiin. Samalla tiedollisten oikeuksien tarkastelu auttaa erilaisia yhteiskunnallisia toimijoita – esimerkiksi julkista hallintoa, oppilaitoksia, mediaa tai tekoälyä käyttöön ottavia yrityksiä – näkemään, miten ne voivat omilla toimillaan vahvistaa ja suojella tiedollisia oikeuksia kehittyvän tekoälyn aikakaudella.



Tiedollisten oikeuksien kehikko rakentuu neljästä oikeudesta:

- 1. Pääsy tietoon:** Kaikilla ihmisillä tulee olla yhdenvertainen pääsy tekoälyn kehitystä ja käyttöä koskevaan olennaiseen ja totuudenmukaiseen tietoon.
- 2. Osaaminen ja valmiudet:** Kaikille ihmisille tulee taata yhdenvertaiset valmiudet ymmärtää tekoälyn kehitykseen sisältyviä mahdollisuuksia ja haasteita. Ihmisten on kyettävä varjelemaan henkilökohtaista tai yhteiskunnan etua tämän ymmärryksen pohjalta.
- 3. Osallistuminen:** Kaikille ihmisille tulee suoda yhdenvertainen mahdollisuus vaikuttaa tekoälykehityksen arvopohjaa, suuntaa ja vaikutuksia koskevaan julkiseen keskusteluun ja päätöksentekoon.
- 4. Ei pakkoa tai painostusta:** Kaikille ihmisille tulee turvata yhdenvertaiset mahdollisuudet muodostaa näkemyksiä kehittyvästä tekoälystä, ilman suurinta tiedollista valtaa käyttävien pakottamista tai manipulointia.

Sammanfattning

Den snabba utvecklingen av artificiell intelligens (AI) under de senaste åren har väckt berättigad oro över maktkoncentrationen inom internationell utveckling av AI till en liten grupp företag som innehar de största teknologiska och ekonomiska resurserna. Rädslan är att dessa företag kan bestämma riktningen för utvecklingen av AI utan att rådfråga andra i större utsträckning.

I detta memorandum behandlas AI-utveckling ur perspektivet av kunskapsrelaterade maktförhållanden. Koncentrationen av kunskapsrelaterad makt inom utvecklingen av AI är ett lika allvarligt problem för demokratiska samhällen som koncentrationen av ekonomisk och teknologisk makt. Detta beror på att teknologijättarna fattar många beslut när de utvecklar AI, beslut som har långtgående ekologiska, etiska, sociala eller samhällseliga konsekvenser.

För att dessa beslut i framtiden ska fattas mer transparent och med hänsyn till samhällets helhetsintresse, bör demokratiska samhällen bryta den kunskapsmässiga ojämlikhet som har rotat sig i AI-utvecklingen och stärka medborgarnas AI-aktörskap. Till stöd för detta presenteras i memorandumet en ram för rättigheter relaterade till kunskap, som erbjuder demokratiska samhällen och olika organisationer verktyg för att komma i gång med arbetet. Den ursprungliga ramen har utarbetats av emeritusprofessor i kommunikation Hannu Nieminen, men i detta memorandum har ramen anpassats för att bättre passa AI-utvecklingens kontext.

Ramen för kunskapsrelaterade rättigheter hjälper till att konkretisera på vilket sätt den makt som har ackumulerats hos stora teknologiföretag försvagar andra aktörers förståelse av AI och begränsar medborgarnas möjligheter att påverka de beslut som fattas i samband med utvecklingen av AI. Samtidigt hjälper utvärderingen av kunskapsrättigheter bland olika samhällsaktörer – till exempel offentlig förvaltning, utbildningsinstitutioner, media eller företag som använder AI – att se hur de genom sina egna åtgärder kan stärka och skydda kunskapsrättigheter i en tid av artificiell intelligens.



Ramen för kunskapsrättigheter består av fyra rättigheter:

- 1. Tillgång till information:** Alla människor ska ha lika tillgång till relevant och sanningsenlig information om utveckling och användning av AI.
- 2. Kompetens och färdigheter:** Alla människor ska garanteras lika möjligheter att förstå de möjligheter och utmaningar som utvecklingen av AI innebär. Människor måste kunna skydda personliga eller samhällsintressen baserat på denna förståelse.
- 3. Deltagande:** Alla människor ska ges lika möjlighet att påverka den offentliga diskussionen och beslutsfattandet om AI-utvecklingens värdegrund, riktning och konsekvenser.
- 4. Inget tvång eller påtryckning:** Alla människor ska garanteras lika möjligheter att bilda sig åsikter om AI, fria från tvång eller manipulation från de som innehar den största makten angående kunskap.

Summary

The rapid development of artificial intelligence (AI) in recent years has raised justified concerns about the concentration of power in international AI development in the hands of a small group of tech companies that possess the largest technological and financial resources. The fear is that these companies can dictate the direction of AI development without consulting others extensively.

This memorandum addresses AI development from the perspective of epistemic, or informational, power. The concentration of epistemic power in AI development is as serious a problem for democratic societies as the concentration of economic and technological power. This is because technology giants make many decisions when developing AI, decisions that have far-reaching ecological, ethical, social, or societal impacts.

To ensure that these decisions are made more transparently in the future and considering the overall benefit of society, democratic societies should dismantle the epistemic inequality that has entrenched itself in AI development and strengthen citizens' AI agency. This memorandum presents a framework for epistemic rights, which offers democratic societies and various organizations ways to get started with the work. The original framework was developed by professor emeritus of communication Hannu Nieminen, but in this memorandum, the framework has been adapted to better fit the context of AI development.

The framework for epistemic rights helps to concretely illustrate how the epistemic power accumulated by large tech companies weakens the AI understanding of other societal actors and narrows citizens' opportunities to influence the decisions made in connection with AI development. At the same time, the examination of epistemic rights helps various societal actors – such as public administration, educational institutions, media, or companies adopting AI – to see how they can strengthen and protect epistemic rights in the emerging AI era through their own actions.



The framework for epistemic rights consists of four rights:

- 1. Access to information:** All people should have equal access to relevant and truthful information about AI development and use.
- 2. Competence and skills:** All people should be guaranteed equal opportunities to understand the possibilities and challenges involved in AI development. People must be able to protect personal or societal interests based on this understanding.
- 3. Participation:** All people should be given equal opportunities to influence the public discussion and decision-making regarding the value base, direction, and impacts of AI development.
- 4. No coercion or pressure:** All people should be guaranteed equal opportunities to form opinions about emerging AI, free from coercion or manipulation by those who hold the greatest epistemic power.

1. Tiedollinen epätasa-arvo on polttava demokratia-kysymys digitalisoituvissa yhteiskunnissa

Digitaalisen teknologian kehityksen ja demokratian välisistä yhteyksistä puhutaan tavallisesti korostaen sitä, miten digitaalinen teknologia on rapauttanut demokratian perustaa. Huomiota kiinnitetään esimerkiksi suosittelualgoritmeihin, jotka ovat muuttaneet mediasisältöjen kulutusta. Yhtä tuttuja ovat pohdinnat siitä, että sosiaalisen median nousu on johtanut perinteisen uutismedian ahdinkoon. Viimeaikaisten tekoälyn kehitysloikkien jälkeen julkisessa keskustelussa on noussut esiin myös huoli tekoälyllä tuotetusta disinformaatiosta ja tekoälyn valjastamisesta vaalivaikuttamisen välineeksi.

Tässä muistiossa demokratiaa ja digitaalisen teknologian kehitystä ei kuitenkaan tarkastella täysin vastakkaisina, erilaisina ilmiöinä. Sen sijaan huomiota kiinnitetään molempia yhdistävään keskeiseen haasteeseen: sekä demokraattiselle päätöksenteolle että digitaalisen teknologian kehittämiseen on tyypillistä, että ne synnyttävät tiedollista epätasa-arvoa.

Molemmissa tapauksissa punnittujen päätösten tekeminen edellyttää laaja-alaista ymmärrystä aiheesta sekä perehtymistä monenlaisiin erityiskysymyksiin ja yksityiskohtiin. Käytännössä syvälinen tietämys kasautuu usein pienehkölle joukolle, jota tässä tekstissä kutsutaan tiedolliseksi eliitiksi. Tällä ryhmällä on hallussaan merkityksellistä tietoa ja osaamista, jonka pohjalta muodostetaan tilannekuvaa, punnitaan vaihtoehtoisia etene- missuuntia ja tehdään päätöksiä. Suurella osalla kansalaisista tällaista tietä- mystä ei ole. Asetelma synnyttää syviä tiedollisia kuiluja eliitin ja muiden välille. Kun näin jatkuu pitkään, kansalaisten kokemus siitä, että heidän näkemyksiään ei kuulla tai huomioda, voi alkaa murentaa luottamusta päätöksentekoon ja nakertaa tehtyjen päätösten oikeutusta.

Tiedollisen epätasa-arvon ongelmat demokraattisessa päätöksenteossa on nostettu aikaisemminkin esiin. Suomessa on pohdittu esimerkiksi sitä, onko valtionhallinnon virkahenkilöillä tai muilla asiantuntijoilla, joilla on käytössään paljon päätöksentekoon vaikuttavaa, arvokasta tietoa, liikaa vaikutusvaltaa poliittiseen päätöksentekoon. On puhuttu lisääntyvästä asiantuntijavallasta ja siitä, miten asiantuntijoiden käyttämä kieli sekä heidän suosimansa tiedon tuottamisen ja esittämisen tavat estävät muita osal- listumasta tiedon keruuseen ja tulkintaan (Ylöstalo 2020).

Digitaalisten teknologioiden kehittämisen yhteydessä esiintyvään tiedolliseen epätasa-arvoon on sen sijaan kiinnitetty vain vähän huomiota. Äkkipäätä väite siitä, että tiedolliset kuilut muodostaisivat ongelman digitaalisen teknologian kehittämisessä saattaakin kuulostaakin kummalliselta. Miksi tiedollisesta epätasa-arvosta pitäisi ylipäättään huolestua aikana, jolloin tietoa eri muodoissaan on tarjolla enemmän kuin koskaan ennen?

Vastaus on kuitenkin yksinkertainen. Tiedollisiin kuiluihin on kiinnitettävä huomiota, koska siirtyminen nykyaikaiseen, verkottuneeseen tietoyhteiskuntaan ei ole poistanut tiedollista epätasa-arvoa. Itse asiassa suuntaus digitaalisten teknologioiden kehittämisessä näyttää olleen miltei päinvastainen. Viestintä- ja tietoteknologian nopea kehitys on kyllä tuonut entistä suuremmat kansanjoukot nykyaikaisen tiedonvälityksen sekä edistyneiden tiedonkäsittelyn ja -tuotannon työkalujen piiriin. Samalla muutos on tehnyt digitaalisesta media- ja informaatioympäristöstä entistä epävarmammaksi sekä luonut voimakkaita teknologiajättejä, jotka kykenevät määräämään digitaalisten teknologioiden kehityksen tahdin. Suurimmat teknologiayritykset pystyvät keräämään valtavia määriä digitaalisessa muodossa olevaa dataa, josta ne jalostavat liiketoiminnallisesti arvokasta tietoa. Niillä on käytössään ylivoimaiset tutkimus- ja tuotekehitysresurssit. Näin taloudellinen, teknologinen ja tiedollinen valta keskittyvät näille jätteille. Samalla niille muodostuu edellytykset ohjata yhteiskunnallisen keskustelun agenda ja sävyä.

Digitalisoituville yhteiskunnille yllä kuvattu kehityskulku on luonut uuden merkittävän demokratiahaasteen. Suomalaisen ja eurooppalaisen demokratian yksi kulmakivi on kansalaisille laissakin taattu oikeus vaikuttaa heitä itseään koskeviin päätöksiin. Oikeuden tulisi ulottua myös digitaalisen teknologian kehittämisessä tehtäviin valintoihin – varsinkin silloin, kun teknologiakehitystä koskevat päätökset vaikuttavat merkittävästi ihmisten arkeen tai yhteiskunnan perusrakenteisiin.

Digitaaliseen teknologiakehitykseen pesiytynyt tiedollinen epätasa-arvo uhkaa kuitenkin vesittää tämän oikeuden. Kun kansalaisilla – tutkijoilla, yhteiskunnallisilla päättäjillä, kansalaisyhteiskunnan aktiiveilla tai digitaalisten palveluiden käyttäjillä – ei ole riittävää pääsyä tärkeiden digitaalisten teknologioiden kehittämistä koskevaan olennaiseen tietoon tai heiltä puuttuu kykyä tulkita saatavilla olevaa tietoa, puntarointi teknologiakehityksen yhteydessä tehtävistä tärkeistä päätöksistä jää vajavaiseksi. Samalla kasvaa riski, että teknologiayritykset tekevät palveluita kehittäessään liian lyhytnäköisiä päätöksiä eivätkä huomioi tarpeeksi esimerkiksi pidemmän aikavälin ekologisia, sosiaalisia tai yhteiskunnallisia seurauksia.

Tämän muistion seuraavissa luvuissa aihepiiriin pureudutaan syvemälle. Tavoitteena on osoittaa, millaiset mekanismit synnyttävät ja ylläpitävät tiedollista epätasa-arvoa digitaalisten teknologioiden kehittämisessä. Samalla etsitään keinoja tiedollisten kuilujen purkamiseksi.

Kuva 1. Tiedolliset oikeudet ihmisten tiedollisen toimintakyvyn vahvistajina**PÄÄSY TIETOON**

Ihmisille tulee taata yhdenvertainen pääsy relevanttiin ja totuudenmukaiseen tietoon.

OSAAMINEN, VALMIUDET

Ihmisille tulee taata yhdenvertainen mahdollisuus kehittää omia valmiuksiaan arvioida ja soveltaa tietoa sekä omaksi että yleiseksi hyödyksi.

OSALLISTUMINEN

Ihmisille tulee taata yhdenvertaiset mahdollisuudet osallistua julkiseen keskusteluun, jossa tehdään yleistä etua koskevia päätöksiä tai muodostetaan niitä koskevia kantoja.

EI PAKKOA TAI PAINOSTUSTA

Ihmisille tulee taata päätösten teossa yhtäläinen vapaus ulkopuolisesta vaikuttamisesta ja painostuksesta.

Tiedollisten oikeuksien kehikossa keskeiset kansalaisten tiedollista toimijuutta vahvistavat toimenpiteet on tiivistetty neljään tiedolliseen oikeuteen.

Lähde: Nieminen 2024.

Lähtökohdan analyysille muodostaa viestinnän emeritusprofessori Hannu Niemisen (2024) hiljattain julkaisema tiedollisten oikeuksien kehikko (ks. Kuva 1). Tiedollisten oikeuksien kehikon taustalla on näkemys, että nykyaikaisissa demokratioissa tieto ja ymmärrys tulisi nähdä perustavanlaatuisina kansalaisoikeuksina, joita ilman demokraattiset yhteiskunnat eivät pysty takaamaan ihmisille yhdenvertaisia mahdollisuuksia osallistua päätöksentekoon. Jotta toimivan demokratian perustaa voidaan vahvistaa, demokraattisten yhteiskuntien ja keskeisten yhteiskunnallisten instituutioiden, kuten median, julkisen hallinnon ja oppilaitosten, pitäisi parantaa ihmisten valmiuksia arvioida ja käsitellä tarjottua tietoa sekä vahvistaa yleisiä mahdollisuuksia osallistua yhteiseen mielipiteenmuodostukseen yhteiskunnallisesti tärkeistä asioista.

Tässä muistiossa tiedollisten oikeuksien kehikkoa sovelletaan uusien digitaalisten teknologioiden kehittämiseen ja käyttöönottoon liittyviin ongelmiin. Tavoitteena on auttaa tietoa tuottavia ja hyödyntäviä organisaatioita pohtimaan, miten ne voivat omalta osaltaan vahvistaa yksilöiden tiedollisia oikeuksia nopeasti etenevän digitalisaation aikana. Kehikkoa voivat hyödyntää myös erilaiset uutta teknologiaa kehittävät yritykset ja muut organisaatiot, kun ne arvioivat omien tuotteidensa ja ideoidensa ”demokratiakestävyyttä”. Päätäjille tai päätöksiä valmisteleville virkamiehille tiedollisten oikeuksien kehikko voi puolestaan toimia työkaluna nykyistä strategisempaan ja ihmislähtöisempään päätöksentekoon esimerkiksi tilanteissa, joissa otetaan käyttöön tai kehitetään kansalaisille suunnattuja digitaalisia palveluita.

Muistiossa teemaa tarkastellaan ensisijaisesti generatiivisen tekoälyn kehityksen kontekstissa. Viime vuosina koettu generatiivisen tekoälyn kehitysluokka on jo kiihdyttänyt julkista keskustelua siitä, keskittyykö tekoälykehitys liian harvojen yritysten käsiin. Tämä muistio syventää keskustelua tuomalla siihen mukaan tiedollisten oikeuksien näkökulman. Monet tekstissä esitetyt huomiot ja johtopäätökset ovat kuitenkin sovellettavissa myös muuhun yhteiskunnallisesti merkittävän teknologian kehitykseen, jossa teknologian kehitystä ohjaa harvalukuinen joukko vaikutusvaltaisia toimijoita.

TIETOLAATIKKO 1

Generatiivisen tekoälykehityksen peruskäsitteitä

Generatiivisella tekoälyllä viitataan kehittyneisiin kielimalleihin, jotka tuottavat uutta sisältöä kirjoitetun tekstin, äänen, kuvien tai videoiden muodossa.

Kehittyneistä kielimalleista käytetään usein myös termiä ”suuri” tai ”laaja” kielimalli. Nämä mallit on koulutettu valtavilla esimerkkiaineistoilla, käytännössä internetistä louhitulla datalla. Niiden koulutuksessa on hyödynnetty monikerroksisia neuroverkkoja. Käytettyjen koulutusmenetelmien ansiosta kehittyneet kielimallit pystyvät suorittamaan monimutkaisia tehtäviä, jotka vaativat edistynyttä puheen, tekstin tai kuvamateriaalin analysointia ja tulkintaa.

Chattibotti on tietokoneohjelma, joka on suunniteltu keskustelemaan ihmisten kanssa joko puhumalla tai kirjoitetun tekstin välityksellä. Generatiivisen tekoälyn kehityksen myötä chattibotit ovat kehittyneet nopeasti, koska niiden rakentamisessa on alettu hyödyntää uusia laajoja kielimalleja. Kehittyneiden kielimallien ansiosta chattibotit kykenevät nykyisin viestimään käyttäjiensä kanssa aiempaa selvästi virheettömämmin, kielellisesti sujuvasti ja ihmisten välistä viestintää uskottavasti jäljitellen.

Lähteet: Sitran Tulevaisuussanasto, päiväämätön; Tietotekniikan termitalkoot, 2025; Tietotekniikan termitalkoot, 2018.

2. Tiedollinen valta generatiivisen tekoälyn kehittämisessä keskittyy teknologiajäteille

Taloudellisen yhteistyön ja kehityksen järjestö OECD julkaisi vuoden 2024 lopussa tekoälykehityksen riskejä ja hyötyjä tarkastelevan katsauksen (OECD 2024). Järjestö arvioi yhdeksi merkittäväksi riskiksi sen, että valta tekoälykehityksessä keskittyy pienelle joukolle vaikutusvaltaisia teknologiayrityksiä. OECD ei ole näkemyksineen yksin. Samaan seikkaan kiinnitettiin hiljattain huomiota esimerkiksi Maailman talousfoorumin julkaisemassa vuosittaisessa riskikartoituksessa (Torkington 2024).

Näissä ja muissa samankaltaisissa analyyseissä vallan keskittymistä tekoälyn kehityksessä tarkastellaan tavallisesti taloudellisesta ja teknologisesta näkökulmasta. Niissä korostetaan, että huipputeknälymallien kehittämisen vaatima pääoma- ja resurssivaltaisuus uhkaa kasata valtaa jättimäisille teknologiayhtiöille, joilla on suurimmat taloudelliset ja teknologiset muskelit. Vähemmälle huomiolle jää kuitenkin se, että samat tekijät voivat johtaa myös tiedollisen vallan keskittymiseen. Tämä johtuu siitä, että tekoälyn kehittämisessä tiedollinen valta perustuu syvällisen asiantuntemuksen ja tekoälyosaamisen ohella myös tekoälymallien kehittämisessä tarvittavien keskeisten teknologisten ja materiaalistien resurssien hallintaan sekä vahvan markkina-aseman ja syvällisen asiakasymmärryksen taitavaan hyödyntämiseen.

Maailman suurimmilla ja vaikutusvaltaisimmilla teknologiayhtiöillä on kaikilla näillä osa-alueilla selvä etulyöntiasema. Yksinkertaistaen: Näillä yrityksillä on isoimmat kassavarat, joiden ansiosta ne voivat turvata pääsynsä tekoälymallien rakentamisessa tarvittaviin tärkeisiin resursseihin, kuten edistyneeseen siruteknologiaan ja laskentatehoon, dataan sekä huipputeknälyosaamiseen.

Generatiivisen tekoälyn kehityksessä tämä yhtälö tulee selkeimmin näkyviin esimerkiksi monien ChatGPT:n kaltaisten keskustelevien chatti-botti-sovellusten taustalla toimivien, edistyneimpien kielimallien kehittämisessä. Näiden kielimallien kehitystä johtavat muutama generatiivisen tekoälyn buumin käynnistänyt kasvuyritys sekä kourallinen isoja teknologiajättejä. Tässä muistiossa huomio kiinnitetään kuitenkin erityisesti jälkimmäisiin, sillä teknologiajätit ovat osoittaneet kykenevänsä käyttämään tehokkaimmin taloudellista ja teknologista valtaa oman tiedollisen valtansa pönkittämiseen.

Isoimmat teknologiayritykset onnistuvat pysyttelemään kielimallien kehityksen kärjessä, koska niiden valtavat teknologiset ja taloudelliset resurssit suovat niille mahdollisuuden sopeutua nopeimmin toimintakentän muuttuviin vaatimuksiin ja olosuhteisiin. Kyky reagoida tehokkaasti on niiden keskeinen valttikortti, jonka ansiosta ne pystyvät lujittamaan tiedollista valtaansa.

Nykyisen tekoälybuumin alussa, 2020-luvun ensimmäisinä vuosina, tämä kyky auttoi isoimpia teknologiayhtiöitä luovimaan menestyksellisesti verisessä kilpailussa, jota kielimalleja kehittävien tahojen välillä käytiin mallien rakentamisessa tarvittavista kriittisistä resursseista. Tässä vaiheessa generatiivisen tekoälyn kehittämiseen virtasi yhtäkkiä maailmanlaajuisesti valtavasti lisää sijoituksia ja alalle syntyi lukuisia uusia tekoälylaboratorioita, joista osa on sittemmin lopettanut tai liitetty osaksi suurempia kilpailijoitaan. Samaan aikaan entistä edistyksellisempien kielimallien kehittäminen vaati jatkuvasti lisää dataa ja laskentatehoa. Kehityksen seurauksena parhaimpien kielimallien rakentamisessa tarvittavien kaikkien tärkeimpien resurssien, edistyneen siru- ja pilvilaskentateknologian sekä internetistä louhittavan harjoitusdatan, kysyntä kasvoi voimakkaasti, eikä näitä resursseja ollut riittävästi tarjolla. Isot teknologiayhtiöt pystyivät kuitenkin valta-voimien taloudellisten ja teknologisten muskeleidensa avulla turvaamaan kriittisten resurssien saannin ja jatkamaan edistyksellisimpien kielimalliensa kehittämistä. Tämän seurauksena niiden tiedollinen valta generatiivisen tekoälyn kehityksessä vahvistui.

Hyvä esimerkiksi tavasta, jolla teknologiajätit pystyivät pönkittämään tiedollista valtaansa tukeutumalla omiin taloudellisiin ja teknologisiin resursseihinsa, on pilvilaskenta. Hakukone Googlen emoyhtiö Alphabet, Microsoft ja Amazon hallitsevat kahta kolmasosaa maailman pilvipalvelumarkkinoista. Markkina-asemansa ansiosta niiden ei tarvinnut kireimmänkään resurssikilpailun aikana huolehtia omien kielimalliensa kehittämisen hidastumisesta laskentatehon saatavuusongelmien vuoksi.

Samalla markkinajohtajuus pilvipalveluissa soi Alphabetille, Microsoftille ja Amazonille mahdollisuuden käyttää omistamaansa laskentatehoa keinona synnyttää generatiivisen tekoälyn kehitykseen itselleen edullisia riippuvuussuhteita. Tämä näkyi räikeimmin siinä, miten jätit käyttivät laskentatehoa neuvotteluvalttinaan, kun ne solmivat miljardien eurojen arvoisia yhteistyösopimuksia edistyneitä kielimalleja kehittävien pienempien startup-yritysten kanssa. Maksuna sijoituksistaan ne luovuttivat startupeille näiden kipeästi tarvitsemaa laskentatehoa, mutta vaativat vastineeksi sopimuksia, jotka takasivat teknologiajäteille pääsyn moniin pienempien yhtiöiden teknologisiin ja tutkimuksellisiin innovaatioihin (Hallamaa 2024).

Teknologiajättien vahvuudet generatiivisen tekoälynkehityksen seuraavassa vaiheessa

Sittemmin pula edistysellisten kielimallien kehittämisessä tarvittavista resursseista on selvästi helpottanut. Tähän ovat vaikuttaneet muun muassa edistyneimpien kielimallien kehittämiseen liittyvät innovaatiot, jotka ovat tehostaneet kielimallien kehittämisessä tarvittavan datan, sirujen sekä laskentatehon käyttöä (Ks. esim. Ng ym. 2025). Onkin todennäköistä, että edistyneimpien kielimallien kehityksessä ollaan siirtymässä pikkuhiljaa uuteen vaiheeseen. Silloin kehityksen kärjessä kulkevat yhtiöt, jotka pystyvät luomaan edistysellisten kielimallien päälle rakentamistaan kuluttajapalveluista yhä moninaisempiin käyttötarpeisiin mukautuvia yleistyökaluja tai tuomaan generatiivista tekoälyä laajasti eri palveluihin. (Ruokonen 2025)

Tällaisen markkinajohtajuuden saavuttaminen edellyttää osin uudenlaista osaamista ja toisenlaisia kyvykkyyksiä kuin generatiivisen tekoälybuumin alkuvaiheessa. Se vaatii kykyä yhdistää alan huippututkimusta suunnitelmalliseen ja jatkuvaan tuotekehitykseen, jossa päämääränä on rakentaa intuitiivisesti toimivia, teknisesti luotettavia ja kasvaviin kuluttajaodotuksiin onnistuneesti vastaavia palveluja. Se vaatii kyvykkyyttä integroida kehittyntä generatiivista tekoälyä sujuvasti erilaisiin työkulkuihin ja palveluprosesseihin.

Generatiivisen tekoälykilvan tässä vaiheessa suurimpien teknologia-yhtiöiden tiedollista valtaa vahvistaa ainakin kaksi merkittävää seikkaa. Ensinnäkin isot teknologia-yhtiöt, joilla on hallussaan suurimmat taloudelliset ja teknologiset resurssit, pystyvät pönkittämään omaa asemaansa generatiivisen tekoälykehityksen kärjessä kehittämällä omaa tekoälyosaamistaan pitkäjänteisesti, laajasti ja kuluja säästelemättä. Tästä todistavat niiden julkistamat lähivuosien investointisuunnitelmat, joiden mukaan yhtiöt aikovat käyttää satoja miljardeja euroja edistyneimpien kielimallien ja niitä tukevan infrastruktuurin kehittämiseen (Bloomberg Intelligence 2024). Samasta asiasta kertoo myös silkka suurten teknologia-yhtiöiden palkkalistoilla työskentelevien tekoälykehittäjien lukumäärä. Erään arvion mukaan pelkästään Amazonilla, Alphabetilla sekä muun muassa Facebookin omistavalla Metalla työskentelee tänä päivänä yhteensä yli 33 000 tekoälyasiantuntijaa (Olson 2023).

Samat teknologiajätit kykenevät myös kartuttamaan tekoäly-ymmärrystään harjoittamalla vilkasta tutkimusyhteistyötä yliopistojen kanssa. Kaikki nämä yhtiöt tekevät monenlaista yhteistyötä tekoälyä tutkivien akateemisten tutkimusryhmien kanssa, eikä tätä kehitystä ole katsottu tutkimusyhteisössä yksinomaan hyvällä. Yhteistyön on pelätty keskittävän kärkitason tekoälytutkimusta entistä pienemmälle joukolle suuria teknologia-yhtiöitä ja huippuyliopistoja (Khanal, Zhang ja Taeihagh 2024). Samalla on kannettu huolta siitä, kasvaako teknologiajättien vaikutusvalta

tiedeyhteisössä liialliseksi (Färber ja Tampakis 2023; Abdalla ja Abdalla 2020; Ahmed, Wahed ja Thompson 2023).

Laaja-alaisen tekoälyosaamisen ohella suurimpien teknologiyhtiöiden tiedollisen vallan toiseksi peruskiveksi generatiivisen tekoälykilvan uudessa vaiheessa vaikuttavat muodostuvan suuret digitaalisten palveluiden portfoliot, joita yhtiöt ovat rakentaneet jo pitkään ennen generatiivisen tekoälyn läpimurtoa. Monilla teknologiajättien omistamilla digitaalisilla palveluilla, kuten hakukoneilla, verkkokaupoilla ja sosiaalisen median palveluilla, on jo entuudestaan valtavat käyttäjämäärät. Kun yhtiöt tuovat palveluihin generatiivista tekoälyä, kertyy niille koko ajan suuria määriä reaaliaikaista, arvokasta dataa generatiivisen tekoälyn käytöstä.

Data auttaa teknologiajättejä selvittämään generatiiviseen tekoälyyn kohdistuvia käyttäjien odotuksia ja uusia tarpeita sekä havaitsemaan mahdollisuuksia laajentaa generatiivisen tekoälyn tarjontansa uusille, täydentäville markkinoille, esimerkiksi terveydenhuoltoon, koulutukseen tai sodankäyntiin (Landi 2024; Stone 2025; Requiroso 2025). Näin suuret teknologiyritykset pystyvät hyödyntämään maailmanlaajuista digitaalisten palveluiden ekosysteemiään tärkeänä generatiivisen tekoälyn kehityksen ja kasvun moottorina.

3. Tekoälykehityksen tiedollisen epätasa-arvon yhteiskunnalliset seuraukset

Tiedollisen vallan keskittyminen teknologiajäteille muodostaa demokraattisille yhteiskunnille uuden kiperän ongelman, jota generatiivisen tekoälyn nopea kehitys ja yleistymisen alleviivaavat. Kun generatiivinen tekoäly lomittuu yhteiskunnan toimintoihin ja palveluihin, sen kehittämisessä tehtävillä päätöksillä on monenlaisia sosiaalisia, ekologisia tai yhteiskunnallisia seurauksia. Sen vuoksi on yhteiskunnan edun mukaista, että näitä päätöksiä tekee riittävän moninainen joukko, laajasti ja läpinäkyvästi erilaisia näkökantoja punniten.

Ongelmaa voidaan havainnollistaa edistyneiden kielimallien varaan rakennettujen keskustelevien chattibottien avulla. Chattibottien käyttö monenlaisissa asiakaspalveluyhteyksissä yleistyy nyt nopeasti. Suomalaisen yhteiskunnan näkökulmasta muutoksessa ei ole kyse pelkästään siitä, onko asiakaspalvelua mahdollista tehostaa chatti-teknologiaa hyödyntämällä. Kyse on myös siitä, että keskustelevista chatti-käyttöliittymistä muodostuu lähitulevaisuudessa tärkeä portti internettiin ja tietoon samalla, kun ne lomittuvat osaksi erilaisia palvelupolkuja. Onko näiden käyttöliittymien nopea yleistymisen kaikkialla pelkästään tavoiteltava kehitys vai tulisiko siinä ottaa huomioon mahdollisuus, että keskustelevia käyttöliittymiä saatetaan käyttää tietoiseen harhaanjohtamiseen, oikean tiedon pimittämiseen tai henkilökohtaisten tietojen varastamiseen (Vrt. esim. Lu 2025; Montreal AI Ethics Institute 2025)? Tai menetämmekö yhteiskuntana jotain muuta tärkeää, mikäli chatti-teknologian käyttöön liittyviä erilaisia näkökulmia ja arvoja ei punnita harkitusti yhdessä?

Viime aikoina esimerkiksi Google on jo liittännyt chattibotti-tyyppisen käyttöliittymän osaksi hakukonettaan. Meta on tuonut uusia generatiivisen tekoälyn ominaisuuksia Facebookiin, Instagramiin ja WhatsAppiin, ja Apple on kytkenässä generatiivista tekoälyä Siri-assistenttiinsa. Niin ikään keskustelevan tekoälyn hyödyntäminen erilaisissa tietotyössä käytettävissä ohjelmistoissa on yleistynyt nopeasti. Seuraavaksi tulossa ovat uuden sukupolven tekoälyassistentit, jotka kykenevät hoitamaan ihmisten puolesta itsenäisesti monenlaisia tehtäviä digitaalisissa ympäristöissä.

Jatkuvasti kasvava tutkimusnäyttö kuitenkin kertoo, että nykyisenkaltaiset generatiivisen tekoälyn kielimallit ovat yhä monin eri tavoin alttiita tuottamaan virheellistä tai väritynyttä tietoa. Tutkimuksissa on jo osoitettu, että puutteelliset kielimallien koulutusmenetelmät voivat jättää

malleihin arvovinoumia, jotka vääristävät niiden päälle rakennettujen tekoälypalveluiden antamia vastauksia (Buschek ja Thorp, päiväämätön). Samoin on todettu, että kielimallit toisinaan hallusinoivat, eli tuottavat täysin keksittyjä vastauksia. Myös niiden toimintalogiikka itsessään saattaa ohjata kielimallit harhaanjohtamaan hienovaraisesti käyttäjiään. (Wong 2024.)

Markkinoille on myös tullut virtuaalista ystävää, rakastettua tai seksikumppania tarjoavia chattibotti-palveluita. Näiden palveluiden käytön on huomattu johtavan ääritapauksissa tunneriippuvuuteen, pahimmillaan traagisin seurauksin (El Atillah 2023; Montgomery 2024). Myös niiden tietosuoja on usein äärimmäisen heikko ja ansaintamallit saalistavia (Caltrider, Rykov ja MacDonald 2024; Kim 2024).

EU:n tekoälyasetus asettaa chattibotti-palveluiden tarjoajille jonkin verran uusia velvoitteita, mutta kaikkiin yhteiskunnallisesti merkittäviin huolenaiheisiin ei lainsäädäntökään voi vastata. Yhteiskuntien on löydettävä muita tapoja reagoida muutoksiin, joita esimerkiksi chattibotti-ystävien ja -kumppanien suosion raju lisääntyminen saattaa jatkossa aiheuttaa vaikkapa sosiaaliin suhteisiimme (Ks. esim. Muldoon 2024; Hill 2025).

Chatti-teknologiaan liittyvät kysymykset ovat kuitenkin vain yksi esimerkki monista generatiivisen tekoälyn kehittämiseen ja käyttöönottoon liittyvistä monimutkaisista yhteiskunnallisista, eettisistä ja ekologisista kysymyksistä, joista ei ole vielä käyty laajaa julkista keskustelua tai yhteistä puntarointia, vaikka generatiivisen tekoälyn kehitys etenee nyt nopein harppauksin. Jotta generatiivista tekoälyä kehitettäisiin tulevana vuosina tasapainoisemmin ja moninaisempia näkökohtia huomioiden, demokraattisten yhteiskuntien tulisikin nyt alkaa määrätietoisesti luoda uusia toimintatapoja ja käytäntöjä, jotka tasoittavat generatiivisen tekoälyn kehittämiseen pesiytynyttä tiedollista epätasa-arvoa.

Keskeisenä ohjenuorana näiden käytäntöjen kehittämisessä tulisi olla, että suurinta tiedollista valtaa generatiivisen tekoälyn kehityksessä käyttävien isojen teknologiayritysten näkökulmat tai arvot eivät saa painottua julkisessa keskustelussa liikaa eikä päätöksiä tämän teknologian kehityssunnista pidä jättää liiaksi niiden käsiin.

Oikea aika uusien toimintamallien kehittämiselle olisi nyt, kun monet kielimallien kehittämiseen liittyvät normit, arvot, standardit, toimintamallit – ja osaksi jopa niiden kehittämistä ja käyttöä koskeva kansainvälinen lainsäädäntö – ovat vasta muotoutumassa.

TIETOLAATIKKO 2

Tiedollisen vallan keskittyminen tuottaa uusia yhteiskunnallisia, eettisiä ja ekologisia haasteita

Kun suurinta tiedollista valtaa käyttävät teknologiayhtiöt kehittävät generatiivista tekoälyä, ne voivat tehdä melko itsenäisesti päätöksiä, joilla on kauaskantoisia vaikutuksia ihmisten arkeen tai yhteiskuntaan.

Tekijänoikeudet. Edistyneiden kielimallien rakentaminen vaatii valtavia määriä harjoitusdataa, jota louhitaan internetistä. Tämä aineiston on osoitettu sisältävän muun muassa tekijänoikeuden alaista materiaalia ja toisinaan jopa yksityishenkilöiden arkaluonteisia tietoja, kuten henkilökoh-
taisia terveystietoja (King ja Meinhardt 2024; Leffer 2023). EU:n tekoäly-
asetus vaatii edistyksellisten kielimallien kehittäjiä ilmoittamaan nykyistä
läpinäkyvämmiin, mitä aineistoa niiden käyttämä kielimallien harjoitusdata
sisältää. Tekijänoikeuden alaisten materiaalin käytön osalta kansainväliset
oikeustulkinnat siitä, missä rajoissa kielimallien kehittäjillä on oikeus
käyttää tällaista materiaalia, ovat kuitenkin vielä muotoutumatta. Edisty-
neimpiä kielimalleja kehittäviä yrityksiä vastaan on vireillä useita merkittä-
viä, tekijänoikeutta käsitteleviä oikeusjuttuja (Ks. esim. Allyn 2025).

Harjoitusdatan sisältämät vinoumat ja syrjivät asenteet. Edistyneiden kielimallien rakentamisessa käytetty internetistä peräisin oleva harjoitus-
data heijastelee yhteiskuntiin pesiintyneitä vinoutuneita ja syrjiviä asen-
teita. Käytössä olevat kielimallien koulutusmenetelmät eivät pysty koko-
naan estämään näiden asenteiden valumista kielimalleihin ja edelleen
niiden varaan rakennettuihin palveluihin. (Ks. esim. Hoffman ym. 2024).
Tämä tulisi ottaa huomioon sekä arvioitaessa näiden kielimallien tuotta-
man sisällön luotettavuutta että pohdittaessa edistyksellisiä kielimalleja
hyödyntävien palveluiden käyttöönottoa uusissa yhteyksissä.

Chattibottien käytön haasteet. Generatiivista tekoälyä hyödyntävän
chatti-tekniikan kehitykseen asetetaan paljon toiveita, ja chattibottien
käyttö erilaisissa yhteyksissä yleistyy nopeasti. Chat-perusteisia palveluita
kehitetään ja kokeillaan aktiivisesti jo esimerkiksi mielenterveys- ja
koulutussektoreilla (Taleva 2025; Elliott 2024; Chadwick 2025). Koska ala
on vielä nuori ja kehittyy hyvin nopeasti, lisää tutkimustietoa ja ymmärrystä
tarvittaisiin kuitenkin esimerkiksi käyttäjien suhteen muuttumisesta
chat-käyttöliittymiin pitkän ajan kuluessa. Tähänastisissa tutkimuksissa on
jo saatu viitteitä muun muassa siitä, että pitkäaikaisempi chattibottien
käyttö oppimisen tukena ei ehkä paranna koululaisten oppimistuloksia,
vaikka lyhyellä aikavälillä niin usein käykin. (Ks. esim. Araujo ja Bol 2024;
Wu ja Yu 2023.) Tuoreissa tutkimuksissa on myös osoitettu, että nykyisillä
chattiboteilla on yhä vakavia rajoitteita terapiakäytössä (Ks. esim. Moore
ym. 2025).



Tekoälyn valtava energiankulutus ja muut ympäristövaikutukset. Kehittyneimpien kielimallien rakentaminen ja niiden varaan rakennettujen palveluiden tarjoaminen kuluttavat valtavia määriä laskentatehoa ja sähköä. Datakeskusten jäähdytys vaatii myös erittäin paljon vettä. Generatiivisen tekoälyn kasvavaan ekologiseen jalanjälkeen on herätty pikkuhiljaa, mutta alan hiilidioksidipäästöjen ja luontokuormituksen hillitsemisestä tulisi silti käydä aktiivisempaa yhteiskunnallista keskustelua ja etsiä tapoja huonon kehityksen kääntämiseksi. Samaan aikaan Suomessakin tulisi selvittää tarkemmin muun muassa sitä, voiko globaalin tekoälykehityksen synnyttämä datakeskusten rakennusbuumi tuoda mukanaan esimerkiksi alueelliseen energian saatavuuteen liittyviä riskejä. Onko mahdollista, että uuden datakeskuksen sijoittuminen alueelle johtaa paikallisiin energian saatavuusongelmiin suurten energiankulutuspiikkien aikana? Mikäli näin kävisi, se saattaisi pahimmillaan estää muiden paljon energiaa vaativien investointiprojektien tulon alueelle, joilla olisi suurempi vaikutus esimerkiksi paikalliseen työllisyyteen tai talouteen. (Hanhivaara, 2025; Velkova 2024.) Keskustelua tulisi käydä myös siitä, millaisiin käyttökohteisiin generatiivista tekoälyä ylipäänsä kannattaa käyttää. Onko fiksumaa käyttää luonnonvaroja kuluttavaa tekoälyä esimerkiksi uusien tuotteiden markkinointiin?

Kehittyneiden kielimallien rikollinen ja pahatahtoinen käyttö. Kehittyneimpiä kielimalleja ja niiden varaan rakennettuja palveluita voidaan hyödyntää myös pahantahtoisin tai rikollisiin tarkoituksiin, pyrkimyksenä horjuttaa yhteiskuntien vakautta, demokratiaa, turvallisuutta tai yksilöiden perusoikeuksia (Ks. esim. NCSC 2024). Tällaista käyttöä ei voida täysin estää, mutta sitä voidaan vaikeuttaa, mikäli edistyneitä kielimalleja rakentavat yritykset jäljittävät ja ehkäisevät systemaattisesti malliensa pahantahtoista käyttöä ja viestivät havainnoistaan aktiivisesti internet-ekosysteemin erilaisille toimijoille (Ks. esim. Open AI 2025). Demokraattiset yhteiskunnat voivat puolestaan kehittää uusia pelisääntöjä ja toimintatapoja kehittyneimpien kielimallien laittomaan käyttöön liittyvien uhkien torjumiseksi sekä pyrkiä suojelemaan erityisesti haavoittuvimpia ihmisryhmiään tekoälyavusteisilta väärinkäytöksiltä.

4. Tiedolliset oikeudet demokraattisemman tekoälykehityksen perustana

Kun uusia toimintamalleja ryhdytään kehittämään, tämän muistion alussa esitelty tiedollisten oikeuksien kehikko voisi tarjota työlle hyvää pohjaa. Kehikko voisi muodostaa erilaisille yhteiskunnallisille toimijoille yhteisen arvoperustaisen lähtökohdan, joka perustuu jaetulle ymmärrykselle siitä, että nopeasti digitalisoituissa yhteiskunnissa ihmisten toimijuutta ja vaikuttamismahdollisuuksia on suojeltava aiempaa ponnekkaammin.

Sen vuoksi tässä luvussa esitellään tiedollisten oikeuksien kehikosta laajennettu versio, joka on sovitettu edistyksellisten kielimallien ja generatiivisen tekoälyn kehittämisen kontekstiin.

Kehikkoa laadittaessa on myös pyritty huomioimaan, että erityyppiset yhteiskunnalliset toimijat voivat parantaa tiedollista tasa-arvoa hyvin erilaisin toimin. Esimerkiksi julkisella hallinnolla, medialla, oppilaitoksilla ja generatiivista tekoälyä kehittäville yrityksillä on yhteiskunnassa erilainen asema ja tehtävä, ja niiden käytössä olevat resurssit vaihtelevat merkittävästi.

Silti kaikilla niillä voi olla oma tärkeä roolinsa tiedollisten oikeuksien puolustamisessa ja vahvistamisessa. Päivitetty kehikko antaa erilaisille toimijoille mahdollisuuden keskittää huomionsa juuri niihin tiedollisiin oikeuksiin, joiden suojelu nivoutuu luontevimmin omaan päivittäiseen toimintaan.

Pääsy tietoon

Kaikilla ihmisillä tulee olla yhdenvertainen pääsy tekoälyn kehitystä ja käyttöä koskevaan olennaiseen ja totuudenmukaiseen tietoon.

Suurimpia kielimalleja ja generatiivista tekoälyä kehittäviä vaikutusvaltaisia teknologiyhtiöitä on arvosteltu siitä, että ne jakavat ulkopuolisille yhä niukasti tietoa monista kielimalliensa kehityksessä tekemistään tärkeistä päätöksistä (Nolan 2023.). Vaikka kielimallien toimintaa ja kehittämistä koskeva raportointi on viime aikoina hieman lisääntynyt, kielimallikehitykseen liittyy yhä seikkoja, joista raportointi on vähäistä ja läpinäkyvyyttä myydyt sitkeässä (Bommasani ym. 2024 a).

EU:n vuonna 2024 voimaan tullut tekoälyasetus parantaa tiedonsaantia. Asetusta on kuitenkin arvosteltu siitä, että edistyneimpiä kielimalleja kehittävilta yrityksiltä vaaditaan julkista raportointia vähemmän kuin esimerkiksi Euroopan Parlamentti oli toivonut lainsäädäntöprosessin aiemmassa vaiheessa. Tekoälyasetus velvoittaa kehittyneimpiä kielimalleja rakentavia yrityksiä tuottamaan tietoa EU:lle ja unionin jäsenvaltioille sekä organisaatioille, jotka rakentavat palveluja näiden yritysten kehittämien kielimallien varaan. Kielimallien päälle rakennettujen palveluiden käyttäjien – kansalaisten – tiedonsaantia asetus pyrkii sen sijaan turvaamaan vain edellyttämällä, että edistyneitä kielimalleja kehittävät yritykset julkistavat ”riittävän yksityiskohtaisen tiivistelmän” mallien kouluttamista varten kerätystä harjoitusdatasta (Bommasani ym. 2024 b).

Läpinäkyvyysraportoinnin ohella kielimalleja kehittävät yritykset voisivat tuottaa lisää tietoa malleistaan ja niiden päälle rakennetuista palveluista muissakin yhteyksissä, huomioiden erilaiset tiedontarpeet ja käyttötilanteet. Tavallisilla palveluiden käyttäjilläkin saattaa olla eri tilanteissa aivan erilaisia syitä haluta tietoa kielimallin toiminnasta. He voivat haluta ymmärtää paremmin, miten malli käyttää ja yhdistelee dataa. He saattavat haluta oppia lisää kielimallien toiminnasta tai arvioida kielimallin päälle rakennetun palvelun tuottaman sisällön oikeellisuutta. (Vrt. esim. Hantula ja Korkman 2022.)

Yleisemmällä tasolla kyse on usein siitä, että tuotettua tietoa tulisi olla riittävästi ja sitä pitäisi tarjota oikeassa muodossa, jotta eri ihmisryhmät pystyisivät arvioimaan, miten paljon kielimalliin voi eri käyttöyhteyksissä ja -tarkoituksissa luottaa (Bansal ym. 2022.). Tavoitteen saavuttaminen edellyttää, että edistyksellisiä kielimalleja kehittävät yhtiöt parantavat tiedonsaantia sekä ottavat käyttöön monipuolisempia ja ihmislähtöisempiä tapoja jakaa tietoa ja tuottaa ymmärrystä kielimallien toiminnasta (Liao ja Wortman Vaughan 2024).

Suurimpia kielimalleja kehittävien yhtiöiden ohella myös organisaatioiden, jotka ottavat käyttöön kielimallien varaan rakennettuja tekoälypalveluita, tulisi luoda uusia käytäntöjä varmistaakseen, että palveluiden käyttäjät ja muut ulkopuoliset saavat tekoälyn käytöstä riittävästi tietoa. Julkiset organisaatiot voisivat näyttää esimerkkiä sitoutumalla julkaisemaan tietoa siitä, miten ja millaisilla periaatteilla ne tekoälyä käyttävät. Tällainen läpinäkyvyyttä edistävä toiminta olisi luontevaa jatkoa uusille velvoitteille, joita EU:n tekoälyasetus joka tapauksessa asettaa julkiselle hallinnolle. Asetuksen mukaan julkisen sektorin toimijoiden on suoritettava perusoikeusarviointi kaikissa tekoälyhankinnoissaan.

Hyviä käytäntöjä julkisia tekoälyjärjestelmiä koskevan tiedon levittämiseen onkin jo olemassa, esimerkkinä Helsingin kaupungin ylläpitämä julkinen tekoälyrekisteri, jossa on tietoa kaikista Helsingin käyttämistä tekoälyjärjestelmistä. Tekoälyjärjestelmiä koskevan tiedon tuottamisessa

erityistä huomiota tulisi kiinnittää siihen, että tarjottu tieto on ymmärrettävää, saavutettavaa ja laajasti saatavilla.

Kansalaisten tiedonsaannin parantamisen ohella tärkeää olisi myös luoda käytäntöjä, jotka turvaavat nykyistä paremmin kansalaisjärjestöjen ja riippumattomien tutkijoiden pääsyn kielimallien kehitystä ja generatiivisen tekoälyn käyttöä koskevaan keskeiseen tietoon. Kielimallien turvallisuutta ja luotettavuutta selvittävässä tutkijayhteisössä on lisääntynyt viime aikoina huoli riippumattoman tutkimuksen tulevaisuudesta. Tämä johtuu siitä, että edistyneitä kielimalleja kehittävät yritykset eivät ole turvanneet riippumattomalle tutkimukselle samanlaisia edellytyksiä kuin esimerkiksi perinteisen ohjelmistokehityksen puolella on tehty (A Safe Harbour for Independent AI Evaluation 2024; Tiku 2024.).

Osaaminen ja valmiudet

Kaikille ihmisille tulee taata yhdenvertaiset valmiudet ymmärtää tekoälyn kehitykseen sisältyviä mahdollisuuksia ja haasteita.

Ihmisten on kyettävä varjelemaan henkilökohtaista tai yhteiskunnan etua tämän ymmärryksen pohjalta.

Generatiivisen tekoälyn yleistyminen asettaa kasvatukselle ja koulutukselle uudistumisen tarpeita. Tutkimukset osoittavat, että oppilailla ja opettajilla on useita virheellisiä käsityksiä tekoälystä (esim. Mertala ja Fagerlund 2024). Myös osaaminen, joka liittyy tekoälyn ilmiönä, sekä tekoälypedagogiikka ovat vielä suurimmalle osalle kasvatus- ja opetusalan ammattilaisista vierasta. Tutkimuksen mukaan kolme neljästä esi- ja perusopetuksen opettajasta ei ole käyttänyt tekoälysovelluksia vielä lainkaan (Tanhua-Piironen ym. 2025).

Suomen kansallisen digitalisaatiostrategian, digikompassin, keskeisenä tavoitteena on, että Suomi on digitaalisesti sivistynyt ja osaava maa vuonna 2030. Suomi haluaa olla jatkossakin maa, jossa kansalaisilla on tietämystä, osaamista ja valmiuksia toimia aktiivisesti sekä ilmaista itseään turvallisesti ja luovasti digitaalisen ajan yhteiskunnassa. Sen vuoksi suomalaisten digitaalisen sivistyksen vaaliminen edellyttää myös vahvaa koululaitoksen panosta. Tekoälylukutaito ja digitaalinen informaatiolukutaito ovat uusia digitaalisen ajan kansalaistaitoja. Digitaalinen informaatiolukutaito on kykyä löytää, tulkita, analysoida, hallita, ymmärtää, luoda ja levittää informaatiota turvallisesti ja asianmukaisesti erilaisissa digitaalisissa ympäristöissä (Faktabaari 2022). Generatiivisen tekoälyn aikakaudella uudenlaisia digitaalisia ympäristöjä ovat muun muassa kehittyneiden kielimallien päälle rakennetut chat-käyttöliittymät ja digitaaliset assistentit.

Koulujen tekoälyopetus on nojautunut tähän saakka vahvasti media- ja teknologiakasvatuksen perinteeseen. Tällöin pyrkimyksenä on ollut ymmärtää tekoälyn toimintalogiikkaa, sovellutusmahdollisuuksia sekä yksilön suhdetta teknologiaan. Tekoälyn yhteiskunnallisia vaikutuksia sekä yhteyttä demokratiaan ja kansalaisten toimijuuteen on sen sijaan tarkasteltu hyvin ohuesti.

Jatkossa koulujen demokratia- ja ihmisoikeuskasvatus voisi tarjota hyvän keinon korostaa tekoälyn ja demokratian välistä suhdetta sekä vahvistaa kansalaisten tiedollista toimintakykyä ja toimijuutta. Tämä vaatisi kuitenkin uudenlaista pedagogista lähestymistapaa ja edellyttäisi koulujen media- ja teknologiakasvatuksen sekä demokratia- ja ihmisoikeuskasvatuksen uudistumista. Tutkimukset osoittavat, että kansalaisten yhteiskunnallinen osallistuminen eriytyy ja kasautuu jo peruskoulusta lähtien (esim. Kestilä-Kekkonen ym. 2022). Lapsille ja nuorille voitaisiin luoda yhdenvertaisemmat mahdollisuudet ymmärtää teknologista kehitystä muistakin kuin teknisistä näkökulmista tuomalla tekoälyn yhteiskunnallisia vaikutuksia käsitteleviä sisältöjä esimerkiksi osaksi yhteiskuntaopin tai katsomusaineiden opetussisältöjä (Juurola ja Korhonen 2025). Samalla parannettaisiin kansalaisten edellytyksiä osallistua tekoälyä koskevaan keskusteluun ja päätöksentekoon.

Esimerkkiä teknologiakasvatuksen sekä demokratia- ja ihmisoikeuskasvatuksen kehittämiseen voitaisiin hakea muun muassa Skotlannista. Maassa käynnistettiin vuonna 2022 alakoululaisille suunnattu kansalaiskasvatuksellisen tekoälyopetuksen koulutusohjelma jatkona aiemmin laaditulle ihmiskeskeistä tekoälyä korostavalle tekoälystrategialle (Children's Parliament 2024).

Skotlannin tavoin myös Suomessa työn media- ja teknologiakasvatuksen ja demokratia- ja ihmisoikeuskasvatuksen uudistamiseksi tulisi perustua yhtenäiseen kansalliseen näkemykseen. Opetushallitus ja opetus- ja kulttuuriministeriö julkaisivat maaliskuussa 2025 tekoälyn käyttöä koskevan aineiston, jonka tavoitteena on tukea varhaiskasvatuksen, opetuksen ja koulutuksen järjestäjiä tekoälyyn liittyvän osaamisen kehittämisessä sekä edistää tekoälyn vastuullista, turvallista ja innovatiivista käyttöä (Opetushallitus 2025). Aineistoon sisältyvissä suosituksissa käsitellään muun muassa tekoälylukutaitoa, kestäväää kehitystä sekä eettisiä periaatteita, mutta tekoälyn yhteys yksilön toimijuuteen ja demokraattisen yhteiskunnan toimintakykyyn jää hyvin vähäiselle huomiolle. Jatkossa tämäntyyppisillä tekoälysuosituksilla voitaisiin kannustaa kasvatus- ja koulutusalojen toimijoita käymään aktiivisempaa yhteiskunnallista vuoropuhelua tekoälykehityksen suunnasta, ja siten edistää kansalaiskasvatuksellista tekoälyopetusta.

Kasvatus- ja koulutussektorin ohella oma roolinsa ihmisten tekoäly-ymmärryksen ja toimijuuden kehittäjänä on myös medialla, jonka tulisi pyrkiä varmistamaan, että kansalaisilla säilyy riittävä ymmärrys

generatiivisen tekoälykehityksen myönteisistä ja kielteisistä yhteiskunnallisista vaikutuksista sekä vaihtoehtoisista kehityssuunnista. Vuonna 2023 esimerkiksi Yhdysvaltojen suurimpiin uutistoimistoihin kuuluva AP lisäsi tyyliohjekirjaansa osion, jossa määritellään, miten tekoälyä käsitellään journalistisesti (Meir 2023). AP:n uusien linjauksien yksi tavoite oli se, ettei tekoälyteknologioille anneta jutuissa sellaista toimijuutta, joka oikeasti kuuluisi näitä teknologioita kehittäville yrityksille ja ihmisille. Median ohella myös muut tekoälystä viestivät organisaatiot voivat tehdä vastaavia sisäisiä linjauksia siitä, millä käsitteillä ja termeillä ne tekoälystä viestivät, jotta ne eivät päädy inhimillistämään tekoälyä tai vahvistamaan virheellisiä käsityksiä tekoälyn kyvyistä (Ks. esim. Robbins ja Bryson 2023; Kambhampati ym. 2025).

Oppilaitosten ja media-alan toimijoiden lisäksi ihmisten toimijuutta ja tekoäly-ymmärrystä voivat tukea myös yritykset ja muut organisaatiot, jotka ottavat käyttöön tekoälyä hyödyntäviä järjestelmiä. Tähän niitä ohjaa jo EU:n tekoälyasetus. Asetus edellyttää, että tekoälyä käyttävien organisaatioiden on varmistettava henkilöstönsä riittävä tekoälylukutaito (Toikkanen ja Toivanen 2025).

Erityisen tärkeää monipuolisen ja kriittisen tekoälylukutaidon vahvistaminen on jatkossa aloilla, joissa tekoälyjärjestelmät lomittuvat tiiviisti esimerkiksi tiedon tuotantoon tai päätöksentekoon (Ks. esim. Nelimarkka, Mervaala ja Laaksonen 2025). Olennaista on myös, että tekoälyä hyödyntävissä organisaatioissa työskentelee riittävästi eri rooleissa toimivia ihmisiä, joilla on valmiudet tarkastella tekoälyjärjestelmien vaikutuksia erilaisista yhteiskunnallisista näkökulmista sekä mahdollisuus vaikuttaa organisaationsa tekoälyä koskeviin päätöksiin. Sen vuoksi tekoälyä käyttöön ottavien organisaatioiden tulisikin kehittää sisäisiä prosesseja ja työtapoja, jotka takaavat, että ne arvioivat tekoälyn käyttöä ja kehittämistä koskevia merkittäviä linjauksiaan monipuolisesti, myös mahdolliset sosiaaliset, ekologiset tai yhteiskunnalliset vaikutukset huomioiden. Eräs hyödyllinen apuväline erilaisten näkökulmien käsittelyyn saattaisi olla kansainvälinen DEDA-arviointityökalu, jonka avulla organisaation eri toimintoja ja osaamista edustavat ihmiset voivat yhdessä pohtia dataprojektien eettisiä jännitteitä ja valintoja. Työkalusta on tehty hiljattain suomenkielinen versio (Repair-työryhmä 2023).

Työntekijöiden tekoälyvalmiuksien kehittäminen ei kuitenkaan poista organisaatioiden vastuuta kehittää eettisesti kestäviä ja kaikenlaiset ihmisryhmät huomioivia tekoälyratkaisuja. Vastuu tällaisten tekoälyjärjestelmien kehittämisestä sekä turvallisesta käyttöön otosta lankeaa aina järjestelmän kehittäjälle ja sen käyttöön ottaneelle organisaatiolle. Se ei saa olla yksittäisten työntekijöiden valvettuneisuuden varassa.

Osallistuminen

Kaikille ihmisille tulee suoda yhdenvertainen mahdollisuus vaikuttaa tekoälykehityksen arvopohjaa, suuntaa ja vaikutuksia koskevaan julkiseen keskusteluun ja päätöksentekoon.

Teknologia-alalla kärsitään yleisesti työntekijäkunnan liian vähäisestä monimuotoisuudesta, eikä suurimpien kielimallien kehitys tee tästä poikkeusta (Ville 2023.) Esimerkiksi suurten kielimallien kouluttamiseen kuolvassa tärkeässä vahvistusoppimisen työvaiheessa valtavan kokoista mallia kouluttaa usein vain hyvin pieni, taustoiltaan ja ikäjakaumaltaan yhdenmukainen joukko (Buschek ja Thorp, päiväämätön). On osoitettu, että tämän joukon asenteet voivat vaikuttaa kielimallien saamaan koulutukseen, minkä seurauksena malleihin saattaa syntyä esimerkiksi uskonnollisia tai poliittisia vinoumia (Perez ym. 2023.)

Vahvistusoppimisen työvaiheen ohella edistyneimpien kielimallien kehittämisessä on muitakin vaiheita, joissa malleja kehittävien ihmisten valinnoilla on merkittävä vaikutus lopputulokseen. Näitä ovat esimerkiksi harjoitusdatan valinta ja datan puhdistaminen sekä kielimallin julkaisua edeltävä red teaming -vaihe, jolloin mallin käyttäytymisestä pyritään löytämään mahdollisia ongelmia.

Samanaikaisesti edistyneimpiä kielimalleja kehittävät yritykset kärsivät tekoälykehitykseen yleisemminkin liittyvästä haasteesta, vinoutuneen datan ongelmasta. Internetistä peräisin oleva kielimallien harjoitusdata on aina jossain määrin vinoutunutta. Sen seurauksena myös valmis malli uusintaa yhteiskuntien seksistisiä tai rasistisia asenteita, jollei ongelmaan havahduta tai riittävästi puututa mallia kehitettäessä (Ks. Esim.Chatterjee 2023; Valonen 2022).

Pientääkseen riskiä, että edistyneet kielimallit tai niitä hyödyntävät digitaaliset palvelut alkavat vahingossa toisintaa räikeästi värittyneitä näkökulmia tai rasistisia tai seksistisiä asenteita, kielimalleja kehittävien yritysten tulisi varmistaa, että kaikissa tärkeissä työvaiheissa mallia on kehittämässä ja arvioimassa kyllin monimuotoinen joukko. Myös kansalaisyhteiskunnan toimijoiden ja kansalaisten näkökulmien ja huolenaiheiden huomioiminen näissä työvaiheissa olisi tärkeää. Sen vuoksi edistyneitä kielimalleja kehittävien yritysten työkalupakkiin voisivat jatkossa kuulua muun muassa puntaroivat kansalaispaneelit sekä toisia kunnioittavan ja kuuntelevan kansalaiskeskustelun välineiksi kehitetyt, uudenlaiset digitaaliset osallistumisalustat. Nämä niin sanotun osallistavan demokratian menetelmät on kehitetty tasoittamaan yhteiskunnallisen osallistumisen kuiluja. Menetelmien käyttöä on jo kokeiltu pistemäisesti suurimpien kielimallien kehittämisessä, mutta jatkossa niiden käyttö tulisi vakiinnuttaa osaksi generatiivisen tekoälyn kehitystä sekä varmistaa, ettei menetelmiä käytetä pelkkään näennäisosallistamiseen (Anthropic 2023.).

Suurimpia kielimalleja kehittävien yritysten lisäksi kansalaisten osallistumis- ja vaikuttamismahdollisuuksia tekoälyn kehityksessä voisivat vahvistaa etenkin poliittiset päättäjät ja julkinen hallinto. Nämä tahot voisivat edellyttää, että kansalaiset pääsevät vaikuttamaan yleisiin, esimerkiksi kansallisiin tai toimialakohtaisiin visioihin, joita laaditaan tekoälyn käytöstä tai toivotuista kehityssuunnista. Kansalaisilla tulisi myös olla mahdollisuus vaikuttaa linjauksiin, joilla ohjataan ja hallitaan tekoälykehityksen aiheuttamia yhteiskunnallisia muutoksia.

Keinoja tällaiseen kansalaisten osallistumiseen tarjoaisivat osin samat osallistuvan demokratian menetelmät, joita edistyneitä kielimalleja kehittävät yrityksetkin voisivat hyödyntää. Esimerkiksi Belgiassa on jo järjestetty puntaroiva kansalaispaneeli, joka tuotti suosituksia tekoälykehityksen tärkeimmistä yhteiskunnallisista näkökulmista. Yhteisten kantojen muodostusta paneelissa tuki joukko asiantuntijoita, joiden tehtävänä oli varmistaa, että kaikilla belgialaispanelisteilla oli riittävät valmiudet puntaroida monitahoisesti tekoälykehitykseen liittyviä erilaisia näkökulmia, hyötyjä ja riskejä. (Bürgerrat 2024.)

Ei pakkoa tai painostusta

Kaikille ihmisille tulee turvata yhdenvertaiset mahdollisuudet muodostaa näkemyksiä kehittyvästä tekoälystä, ilman suurinta tiedollista valtaa käyttävien pakottamista tai manipulaatiota.

Suurten teknologiayritysten kasvanut lobbausvoima on herättänyt viime vuosina huolta varsinkin kansalaisyhteiskunnan toimijoiden joukossa (Ks. esim. Corporate Europe Observatory 2025 a). Teknologiajättien isot taloudelliset ja teknologiset muskelit takaavat niille mittavat resurssit, joita ne voivat kanavoida systemaattiseen lobbaamiseen. Tästä osoituksena suurimpien teknologiayhtiöiden, kuten Alphabetin, Applen, Facebookin ja Microsoftin, lobbauskulut Euroopan Unionissa kasvoivat vuosien 2021–2023 välillä 16,5 prosentilla 113 miljoonaan euroon (Djurickovic 2023). Käytännössä teknologiayhtiöiden lobbarit tarjoilevat usein esimerkiksi tarkkoja ehdotuksia lainsäädäntöteksteiksi ajaakseen edustamiensa tahojen etuja.

Euroopassa suuret teknologiayritykset ovat pyrkineet viime vuosina erityisesti vaikuttamaan päättäjiin ja viranomaisiin, jotka ovat valmistelleet EU:n tekoälyasetusta ja sen toimeenpanoa (Ks. esim. Corporate Europe Observatory 2025 b). Lainsäädäntötyön lopputulos voi vinoutua, jos suurilla yhtiöillä on toisiin toimijoihin, kuten kansalaisjärjestöihin, verrattuna suhteettoman suuret mahdollisuudet käyttää aikaa, asiantuntemusta ja muita resursseja vaikuttaakseen lainsäädännön valmisteluun. Jos erilaisia tahoja ja näkökulmia ei ehditä kuulla tasapuolisesti tai kaikilla tahoilla ei

ole samanlaisia resursseja osallistumiseen tai vaikuttamiseen, demokraattisen päätöksenteon periaatteet vaarantuvat (Djurickovic 2023).

Kansalaisjärjestöt ovat vaatineet, että yritysten liiallinen edustus EU:n erilaisissa neuvoa-antavissa ryhmissä, lobbaustapaamisissa ja poliittisissa prosesseissa tulisi estää. Ne ja muutkin tahot ovat myös peränneet uudistuksia ja tarkennuksia EU:n avoimuusrekisteriin, joka ei nykyisellään anna niiden mielestä riittävän kattavaa ja yhdenmukaista kuvaa eri tahojen lobbaustoimista EU:n elimissä. (European Court of Auditors 2024; Corporate Europe Observatory 2024.)

Kansalaisjärjestöille tulisi myös turvata riittävät taloudelliset resurssit osallistua lainsäädännön valmisteluun tai lainsäädäntötyötä mahdollisesti seuraavaan tärkeään, teknisten standardien luomisvaiheeseen. Esimerkiksi EU:n tekoälyasetukseen liittyvässä standardointityössä on jo pantu merkille, että kansalaisjärjestöjen osallistumista standardien luomiseen hankaloittavat käytännössä usein taloudelliset esteet. Jotta kansalaisjärjestöt voisivat osallistua nykyistä aktiivisemmin esimerkiksi tekoälystandardien laatimiseen, standardointityöhön osallistumisen tueksi luotuja rahoituskanavia tulisikin yhä kehittää. (Galvagna 2023)

Suoran lobbaamisen ohella jatkossa olisi syytä kiinnittää systemaattisemmin huomiota myös siihen, kuinka isot teknologiayhtiöt saattavat harjoittaa epäsuoempaa vaikuttamista hallussaan olevien tekoälypalveluiden välityksellä. Generatiivisen tekoälyn kehityksen näkökulmasta huolta lähivuosina herättävät esimerkiksi nopeasti kehittyvät tekoälyassistentit, joiden viestinnästä pyritään tietoisesti tekemään ihmismäistä. Tutkimusten perusteella tiedetään jo, että mitä uskottavammin nämä käyttöliittymät jäljittelevät tunnereaktioita tai tuottavat vakuuttavan tuntuista mielipiteitä, sitä todennäköisemmin niiden käyttäjät alkavat mieltää ne tahoiksi, jolla on oma persoona. Ihmisenkaltaisuus saattaa myös johtaa käyttäjät luottamaan chat-käyttöliittymiin sokeammin tai vakuuttaa heidät siitä, että käyttöliittymien kanssa on mahdollista muodostaa vuorovaikutteinen, sosiaalinen suhde. (Cornelius ja Leidner 2021; van Wynsberghe 2022.)

Kehittynyttä chat-teknologiaa hyödyntävien yritysten tai muiden tahojen on mahdollista pyrkiä manipuloimaan chattibottien käyttäjiä tietoisesti, eikä demokraattisten yhteiskuntien tulisi luottaa liian sinisilmäisesti siihen, että näin ei tehdä. Maailmanpoliittisessa tilanteessa, jossa teknologia- ja palvelukehityksessä tehtyjä valintoja saattavat entistä vahvemmin ohjata myös avoimen poliittiset motiivit, ei ole täysin poissuljettua, että ihmisen kaltaisesti viestivää, tunteisiin vetoavaa chat-teknologiaa pyrittäisiin hyödyntämään oman poliittisen tai yhteiskunnallisen agendan edistämiseen (Vrt. Jackson, 2024; Laforge 2025). Demokraattisissa yhteiskunnissa tulisi-kin nyt käydä laajempaa keskustelua siitä, miten varsinkin haavoittuvimmassa asemassa olevia ihmisiä kyetään suojaamaan generatiivisen tekoälyteknologian manipulatiiviselta ja epärehelliseltä käytöltä.

5. Jälkisanat: Kolme askelta tiedollisten oikeuksien soveltamiseen käytännössä

Tässä muistiossa esitetty hahmotelma tekoälyn kehittämiseen liittyvistä tiedollisista oikeuksista on ensimmäinen yritys istuttaa tiedollisten oikeuksien ajattelua teknologiakehityksen kontekstiin. Siihen tulisi siksi suhtautua ensisijaisesti keskustelunavauksena, josta ajattelu voi jatkossa syventyä ja kehittyä.

Jo nykyisellään hahmotelma tarjoaa kuitenkin monenlaisille organisaatioille mahdollisuuden käyttää tiedollisia oikeuksia oman toimintansa kiritäjinä. Kehikon käyttö auttaa varmistumaan, että tiedollisten oikeuksien näkökulmat otetaan huomioon riittävän monipuolisesti, laajasti ja oikea-aikaisesti.

Seuraavat kolme askelta auttavat tiedollisista oikeuksista kiinnostuneita pääsemään työssä alkuun.

1. Liitä tiedolliset oikeudet olemassa oleviin prosesseihin ja toimintoihin

Tarkastele aluksi organisaatiosi prosesseja ja käytäntöjä ja arvioi, missä yhteyksissä tiedollisten oikeuksien kehikon käyttö olisi olennaisinta ja vaikuttavinta. Tiedollisten oikeuksien kehikkoa voidaan käyttää esimerkiksi eettisten periaatteiden tai muiden keskeisten ohjeistusten demokratialäpivalaisuun. Kehikko voi myös tarjota tärkeän näkökulman strategia-työhön tai auttaa arvioimaan, kuinka kestävä jokin uusi, tekoälyä hyödyntävä ratkaisu on demokratian näkökulmasta. Tiedollisten oikeuksien kehikon käyttö tekoälyä hyödyntävien organisaatioiden vastuullisuustyön tukena saattaa puolestaan kirittää pohtimaan, miten voi vahvistaa oman toimintansa positiivisia yhteiskunnallisia vaikutuksia, eli niin sanottua yhteiskunnallista tai sosiaalista kädenjälkeä.

2. Luo kumppanuuksia ennakkoluulottomasti yli totuttujen rajojen

Tekoälykehitys näyttäytyy julkisessa keskustelussa usein lähes väistämättömänä kehityskulkuna, jonka etenemiseen kansalaiset tai suurin osa organisaatioistakaan eivät voi vaikuttaa. Samaan aikaan kehitykseen liittyy ihmisten mielissä epävarmuuksia ja huolia, ja suomalaisten luottamus tekoälyyn on kansainvälisissä vertailuissa alhaisella tasolla (Gillespie ym. 2023).

Tiedollisten oikeuksien kehikko tarjoaa keinon synnyttää tälle kehitykselle vastavoima. Se suo sellaisillekin organisaatioille, jotka eivät perinteisesti näyttäyty teknologian kehityksestä käytävässä julkisessa keskustelussa aktiivisina toimijoina tai muutoksentekijöinä, mahdollisuuden suunnata omia resurssejaan ja voimavarojaan demokraattisemman ja tasa-arvoisemman tekoälykehityksen vahvistamiseen. Samanaikaisesti kehikko luo näille organisaatioille yhteisen arvopohjaisen perustan. Tältä pohjalta on mahdollista solmia keskinäisiä kumppanuuksia laajalla rintamalla, perinteiset toimialarajat ja muut siilot ylittäen. Erilaiset yhteiskunnalliset toimijat voivat myös käyttää tiedollisten oikeuksien kehikkoa yhteisenä pohjana, jonka avulla ne laativat tiekartan tiedollista tasa-arvoa vahvistavista toimista.

3. Tunnista tulenkantajat

Tiedollisten oikeuksien ajattelu ei vakiinnu yhteiskunnassa, jollei se saa tuekseen tulenkantajia, jotka hyödyntävät strategisesti omia vahvuuksiaan tiedollisia oikeuksia levittääkseen. Siksi on tärkeää, että organisaatiot, joiden tiedollisia oikeuksia vahvistavat toimet voivat tuottaa pysyvämpiä ja systemisempiä vaikutuksia, tunnistavat omat vaikutusmahdollisuutensa sekä vipukohdat, joissa toimilla on erityisen suuri vaikutus.

Tiedollisten oikeuksien tulenkantajina voivat toimia esimerkiksi soveltavaa tutkimusta ja kehitystoimintaa tekevät – tai sitä rahoituksellaan tukevat – organisaatiot. Niiden on mahdollista suunnata resurssejaan kehitystyöhön ja kokeiluihin, joissa luodaan eri konteksteihin tai toimialoille soveltuvia, tiedollisia oikeuksia vahvistavia toimintamalleja. Avaintoimijoita voivat olla myös julkisen hallinnon organisaatiot. Ne voivat integroida tiedollisten oikeuksien näkökulmia julkista hallintoa tai tiettyä toimialaa koskeviin ohjeistuksiinsa, suosituksiinsa, politiikkoihinsa tai linjauksiinsa, jotka kytkeytyvät esimerkiksi demokratiatyöhön, tekoälykasvatukseen tai tekoälyn vastuulliseen käyttöön. Julkisen hallinnon on myös mahdollista näyttää esimerkkiä muille toimijoille ottamalla käyttöön tiedollisia oikeuksia vahvistavia käytäntöjä, jotka kehittävät tekoälyn käytön läpinäkyvyyttä tai parantavat kansalaisten mahdollisuuksia vaikuttaa tekoälyn hyödyntämistä koskeviin päätöksiin.

Lähteet

Ahmed, N., Wahed M. ja Thompson N. 2023. The growing influence of industry in AI research. Science 379, 884-886 (2023). (Haettu 4.10.2024)

Allyn, Bobby. 2025. 'The New York Times' takes OpenAI to court. ChatGPT's future could be on the line. NPR, 14.1.2025. (Haettu 12.2.2025)

Anthropic. 2023. Collective Constitutional AI: Aligning a Language Model with Public Input. Anthropic, 17.10.2023. (Haettu 4.10.2024)

Araujo, Theo ja Bol, Nadine. 2024. From speaking like a person to being personal: The effects of personalized, regular interactions with conversational agents. Computers in Human Behavior: Artificial Humans, Volume 2, Issue 1, 2024.

A Safe Harbour for Independent AI Evaluation. 2024. AI-tutkijoiden avoin kirje, maaliskuu 2024.

Bansal ym. 2022. Workshop on Trust and Reliance in AI-Human Teams (TRAIT). CHI EA '22: CHI Conference on Human Factors in Computing Systems Extended Abstracts, artikkeli no. 116, 1–6.

Bloomberg Intelligence. 2024. Big tech 2025 capex may hit \$200 billion as gen-AI demand booms. Bloomberg Professional Services, 4.10.2024.

Bommasani, Rishi ym. 2024 a. The 2024 Foundation Model Transparency Index. Transactions on Machine Learning Research, 12/2024.

Bommasani, Rishi ym. 2024 b. Foundation Models under the EU AI Act. Stanford Center for Research on Foundation Models, 1.8.2024. (Haettu 13.11.2024)

Buschek, C. ja Thorp, J. Päiväamäätön. Models All the Way Down. Knowing Machines. (Haettu 13.9.2024)

Bürgerrat. 2024. Artificial intelligence discussed with wit. 26.5.2024. (Haettu 13.9.2024)

Caltrider, J., Rykov M. ja MacDonald Z. 2024. Romantic AI Chatbots Don't Have Your Privacy at Heart. Mozilla Foundation, 14.2.2024. (Haettu 12.2.2025)

Calvagna, Christine. 2023. Inclusive AI governance. Ada Lovelace Institute, 30.3.2023. (Haettu 4.10.2024)

Chadwick, Lauren. 2025. Estonia announces a new AI initiative with OpenAI to roll out chatbots in schools. Euronews, 25.2.2025. (Haettu 31.3.2025)

Chatterjee, Mohar. 2023. Bias in AI is real. But it doesn't have to exist. Politico, 14.7.2023. (Haettu 31.3.2025)

Children's Parliament. 2024. Exploring Children's Rights and AI. Stage 2 Summary Report.

Cornelius, Samia ja Leidner, Dorothy. 2021. Acceptance of Anthropomorphic Technology: A Literature Review. Proceedings of the 54th Hawaii International Conference on System Sciences (2021) 6422–6431.

Corporate Europe Observatory ym. 2024. Open Letter to EU Commission President Ursula von der Leyen. Corporate Europe Observatory, Lobby Control, Transparency International and The Good Lobby, 15.7.2024. (Haettu 13.11.2024)

Corporate Europe Observatory. 2025 a. Coded for Privileged Access. How Big Tech weakens rules of advanced AI. 30.4.2025. (Haettu 21.5.2025.)

Corporate Europe Observatory. 2025 b. Bias baked in. How Big Tech sets its own AI standards. 9.1.2025. (Haettu 21.5.2025)

Djurickovic, Tamara. 2023. The impact of big tech: Increased lobbying expenses in the EU. Tech.eu, 12.9.2023. (Haettu 13.11.2024)

El Atillah, Imane. 2023. Man ends his life after an AI chatbot 'encouraged' him to sacrifice himself to stop climate change. Euronews, 31.3.2023. (Haettu 19.9.2024)

Elliott, David. 2024. This AI tutor could make humans 10 times smarter, its creator says. World Economic Forum, 29.7.2024. (Haettu 31.3.2025)

European Court of Auditors. 2024. Special report 05/2024: EU Transparency Register – provides useful but limited information on lobbying activities. (Haettu 13.11.2024)

Faktabaari. 2022. Digitaalinen informaatiolukutaito-opas. Toim. Kari Kivinen. Avoin yhteiskunta ry / Faktabaari EDU.

Färber, M., Tampakis, L. 2024. Analyzing the impact of companies on AI research based on publications. Scientometrics 129, 31–63 (2024).

Gillespie, N., Lockey, S., Curtis, C., Pool, J., ja Akbari, A. (2023). Trust in Artificial Intelligence: A Global Study. Queenslandin yliopisto ja KPMG Australia.

Hallamaa, Teemu. 2024. Automaattinen tietojenkäsittely keskittyy valtaviiin datakeskuksiin, mikä synnyttää uudenlaisia portinvartijoita. Yle, julkaistu 27.5.2024. (Haettu 13.11.2024)

Hanhivaara, Jussi. 2025. Asiantuntija kriittinen Suomeen rakennettavista datakeskuksista: Hyöty jää vähäiseksi. Yle, 23.2.2025. (Haettu 28.2.2025)

Hantula, Kirsi ja Korkman, Oskar. 2022. Embracing user unpredictability. Solita

Hill, Kashmir. 2025. They asked ChatGTP questions. The answers sent them spiraling. New York Times, 13.6.2025. (Haettu 15.6.2025)

Hoffman, V. ym. 2024. AI generates covertly racist decisions about people based on their dialect. Nature.com, 28.8.2024. (Haettu 28.2.2025)

Jackson, Dean. An Eight-Year Struggle for Accountability in the US Ends in Defeat. Tech Policy Press, 18.11.2024. (Haettu 6.3.2025)

Juurola, Leenu ja Korhonen, Tiina. 2025. Demokratiakasvatus, lasten oikeudet ja tekoäly: Tilanne- ja tarvekartoitus sekä toimintasuunnitelma. Helsingin yliopiston Innokas-verkoston julkaisuja. (Haettu 20.5.2025)

Kestilä-Kekkonen, Elina ym. 2022. Nuoret kansalaiset. Tilastollinen tutkimus yhdeksäsluokkalaisten kansalaispätevydestä. Oikeusministeriön julkaisuja, Selvityksiä ja ohjeita 2022:11. (Haettu 12.2.2025)

Khambhampati Subbrao ym. 2025. Stop Anthropomorphizing Intermediate Tokens as Reasoning/Thinking Traces! ArXiv-preprint-versio.

Khanal S., Zhang H. ja Taihagh A. 2024. Why and how is the power of Big Tech increasing in the policy process? The case of generative AI, Policy and Society, 2024.

Kim, Andrew. 2024. Is AI Companionship The Next Frontier In Digital Entertainment? ARK Invest, 18.6.2024. (Haettu 12.2.2025)

King, J. ja Meinhart, C. 2024. Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World. Stanford University white paper.

Laforge, Gordon. 2025. Big Tech's Foreign Policy Takeover. Tech Policy Press, 2.5.2025. (Haettu 6.6.2025)

Landi, Heather. 2024. Google's healthcare AI ambitions: New model for personal health coaching and next steps for gen AI. Fierce Healthcare -verkkolehti, 19.3.2024. (Haettu 17.1.2025)

Leffer, Lauren. Your Personal Information Is Probably Being Used to Train Generative AI Models. Scientific American, 19.10.2023.

Liao, Q. Vera ja Wortman Vaughan, Jennifer. 2024. AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. Harvard Data Science Review, 31.5.2024. (Haettu 20.9.2024)

Lu, Donna. 2025. We tried out DeepSeek. It worked well, until we asked it about Tiananmen Square and Taiwan. The Guardian, 28.1.2025. (Haettu 3.2.2025)

Meir, Nicole. 2023. AI guidance, terms added to AP Stylebook. AP, 16.8.2023. (Haettu 20.9.2024)

Mertala, Pekka ja Fagerlund, Janne. 2024. Finnish 5th and 6th graders' misconceptions about artificial intelligence. International Journal of Child-Computer Interaction, 39, Article 100630.

Montgomery, Blake. 2024. Mother says AI chatbot led her son to kill himself in lawsuit against its maker. The Guardian, 23.10.2024. (Haettu 13.11.2024)

Montreal AI Ethics Institute. 2025. The AI Ethics Brief #166: AI Systems in Human Spaces - Accountability and Unintended Consequences. (Haettu 11.6.2025)

Moore, J. ym. 2025. Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers. ArXiv preprint-versio.

Muldoon, James. 2024. 'Maybe we can role-play something fun': When an AI companion wants something more. BBC, 9.10.2024. (Haettu 12.2.2025)

Nelimarkka, Matti; Mervaala, Erkki ja Laaksonen, Sanna-Maaria. 2025. Rajapinta guidelines for research ethics with generative artificial intelligence. Rajapinta ry. – Digitaalisten yhteiskuntatieteiden yhdistys.

Nieminen, Hannu. 2024. Why we need epistemic rights? Teoksessa Epistemic Rights in the Era of Digital Disruption. Toim. Aslama Horowitz ym. Palgrave MacMillan, Cham.

Ng ym. 2025. DeepSeek: The Chinese AI app that has the world talking. BBC, julkaistu 4.2.2025. (Haettu 28.2.2025)

Nolan, Michael. 2023. Llama and ChatGPT Are Not Open-Source. IEEE Spectrum, 27.7.2023. (Haettu 13.11.2024)

NSCS. 2024. The near-term impact of AI on the cyber threat. National Cyber Security Centre, 24.1.2024. (Haettu 12.2.2025)

OECD. 2024. Assessing potential future artificial intelligence risks, benefits and policy imperatives. OECD Artificial Intelligence Papers, No. 27, OECD Publishing, Paris.

Olson, Parmy. 2023. Tech's Armies Are Huge, Yet Struggling to Innovate. Bloomberg, 26.3.2023. (Haettu 13.11.2024)

OpenAI. 2025. Disrupting malicious uses of our models: an update. OpenAI, 21.2.2025.

Opetushallitus. 2025. Tekoäly varhaiskasvatuksessa ja koulutuksessa – lainsäädäntö ja suositukset. (Haettu 20.5.2025)

Perez, Ethan ym. 2023. Discovering Language Model Behaviors with Model-Written Evaluations. Findings of the Association for Computational Linguistics: ACL 2023, pages 13387–13434 July 9-14, 2023.

Repair-työryhmä. 2023. Helsingin yliopiston johtama hanke toi tekoälyn eettisen arvioinnin työkalun Suomeen. Helsingin yliopisto, 25.10.2023. (Haettu 13.11.2024)

Requiroso, Kelvene. 2025. Pentagon Advances Generative AI in Military Operations Amid US-China Tech Race. eWeek, 22.4.2025. (Haettu 20.5.2025)

Robbins, Scott A. ja Bryson, Joanna J. 2023. Do We Collaborate With What We Design? Topics in Cognitive Science 00 (2023) 1–20.

Ruokonen, Mika. 2025. Applying AI in key European industries. Sitra, 27.3.2025.

Sitra. Päiväämätön. Generatiivisen tekoälyn määritelmä Sitran tulevaisuussanastossa.

Stone, Adam. 2025. How Higher Ed Institutions Are Using Built-In Generative AI Tools. EdTech-verkkolehti, 6.2.2025. (Haettu 13.3.2025)

Taleva, Katariina. 2025. Näin mielenterveyden avuksi kehitetty tekoäly toimii – MTV Uutiset testasi. MTV3, 28.1.2025. (Haettu 28.2.2025).

Tanhua-Piironen, Erika ym. 2025. Esi- ja perusopetuksen digitalisaation tilannekuva 2025. Opetus- ja kulttuuriministeriön julkaisuja 2025:13.

Tietotekniikan termitalkoot, 2018. Chattibotin määritelmä. Tietotekniikan termitalkoot, 2018-12-19. Sanastokeskus ry.

Tietotekniikan termitalkoot. 2025. Suuren tai laajan kielimallin määritelmä. Uusia suosituksia Tietotekniikan termitalkoissa 04/2025. Sanastokeskus ry.

Tiku, Nitasha. 2024. Top AI researchers say OpenAI, Meta and more hinder independent evaluations. The Washington Post, 5.3.2024. (Haettu 20.9.2024)

Toikkanen, Tarmo ja Toivanen, Meeri. 2025. Tekoälyasetus tulee – onko lukutaito kunnossa? Sitra, 3.2.2025. (Haettu 13.3.2025)

Torkington, Simon. 2024. These are the 3 biggest emerging risks the world is facing. World Economic Forum. (Haettu 12.12.2024)

Valonen, Tiina. 2022. Tekoälyjärjestelmät täytyy arvioida yhdenvertaisuuden näkökulmasta – tuore arviointityökalu tarjoaa uusia mahdollisuuksia siihen. Yhdenvertaisvaltuutettu, 27.9.2022. (Haettu 16.9.2024)

Velkova, Julia. 2024. Dismantling public values, one data center at the time. Reimagining public values in algorithmic futures -tutkimuskonsortio, 19.2.2024. (Haettu 13.3.2025)

Ville, Lizzie. 2023. System update. Addressing the gender gap in tech. Fawcett Society, 30.10.2023.

van Wynsberghe, Aimee. 2022. Social robots and the risks to reciprocity. AI & SOCIETY (2022) 37:479–485.

Wong, Matteo. 2024. Chatbots are primed to warp reality. The Atlantic, 30.8.2024. (Haettu 12.9.2024)

Wu, Rong ja Yu, Zhonggen. 2023. Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. British Journal of Education Technology, Volume 55, Issue 1, January 2024, 10-33.

Ylöstalo, Hanna. 2020. Asiantuntijavalta politiikassa. Politiikasta-tiedeverkkolehti. (Haettu 24.10.2024)

SITRa

SITRAN MUISTIO KESÄKUU 2025

Sitran muistiot ovat tulevaisuustyömme taustaksi
tuotettuja sisältöjä.

ISBN 978-952-347-419-2 (PDF) www.sitra.fi

SITRA.FI

Itämerenkatu 11-13
PL 160
00181 Helsinki
P. +358 294 618 991
LinkedIn @Sitra