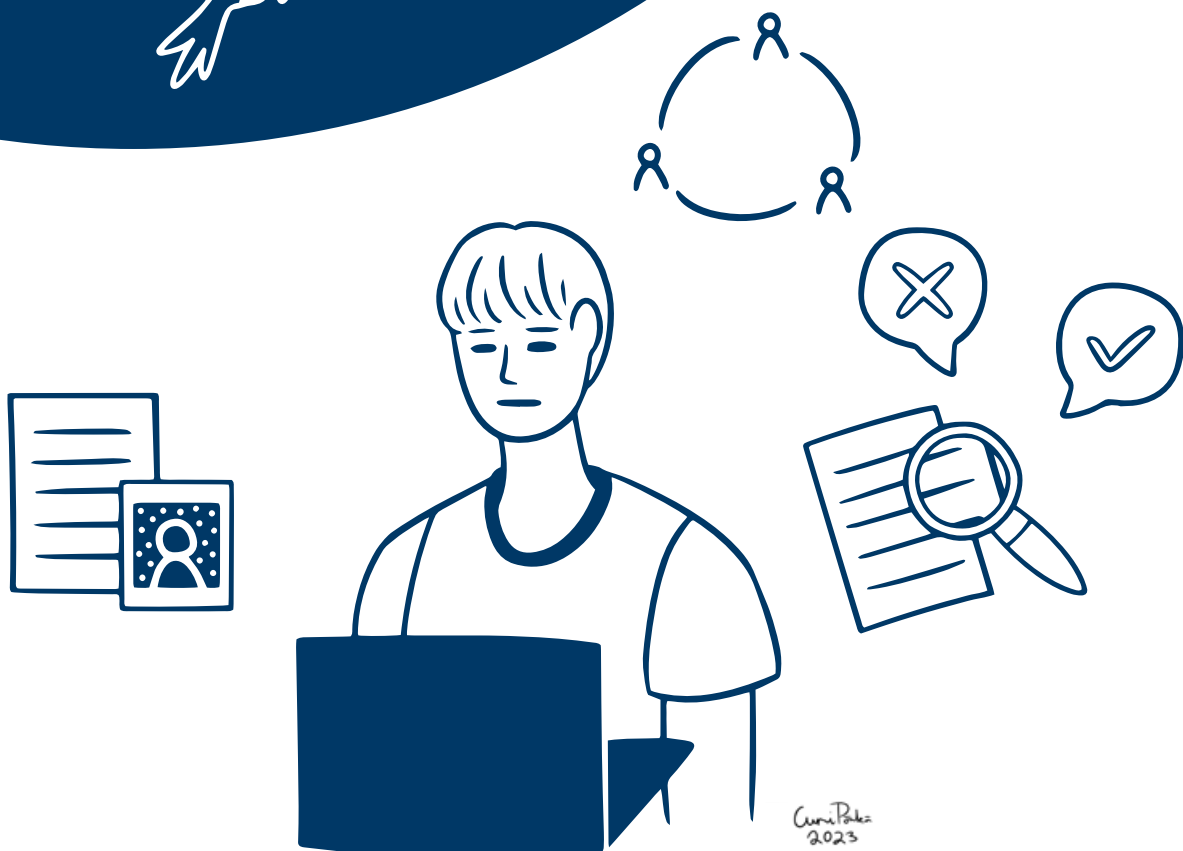


DIGITAALINEN
INFORMAATIOLOKUTAITO

Tekoälyopas



FaktaBaari **EDU**

MINNA ASLAMA HOROWITZ, MATT CRONIN, PIPSA HAVULA, TIINA HÄRKÖNEN, KARI KIVINEN,
HARTO PÖNKÄ, AKI SAARIAHO JA TAPANI SAARINEN
KUVITUS LUMI PÖNKÄ JA DALL-E 3 WWW.BING.COM/CREATE

Digitaalinen informaatiolukutaito: Tekoälyopas

Yleistajuinen opas digitaalisen uutisvirran seuraajille

ISBN 978-952-69148-4-8 (nidottu)

ISBN 978-952-69148-5-5 (PDF)

2023 Avoin yhteiskunta ry / Faktabaari EDU

Vastaava toimittaja: Kari Kivinen

Kirjoittajat: Kari Kivinen, Minna Aslama Horowitz, Matt Cronin, Pipsa Havula, Tiina Härkönen, Harto Pönkä, Aki Saariaho ja Tapani Saarinen.

Kuvitus: Lumi Pönkä ja DALL-E 3

Siteeraukset: Kivinen, K., Aslama Horowitz, M., Cronin, M., Havula, P., Härkönen, T., Pönkä, H., Saariaho, A., Saarinen, T. (2023) Digitaalinen informaatiolukutaito, - Tekoälyopas, Faktabaari, Helsinki

Sisältö

Sanasto	5
Johdanto	6
1. Tekoäly & faktantarkistus	10
Miten keskustelubotit toimivat?	10
Tekoäly ei erota, mikä on totta ja mikä ei	11
Kuka tarkistaisi faktat ja kantaisi vastuun?	12
Syväväärennökset	13
2. Demokratia, osallisuus ja tekoäly	15
EU:n lait suojelevat ihmisiä ja organisaatioita	16
Suomi, päättäjät ja tekoäly	16
Kuka saa vaikuttaa tekoällyn kehitykseen?	16
Haaveena demokraattinen ja inhimillisiin arvoihin perustuva tekoäly	17
3. Tekoälyohjatut sosiaalisen median botit informaatiovaikuttamisessa	18
Mitä eroa on trollilla ja botilla?	18
Bottien tunnistaminen vaatii perusteellista analyysiä	19
Miten generatiivinen tekoäly muuttaa sosiaalisen median botteja?	20
Trollien ja bottien tutkijat ovat ahtaalla	21
Kuinka yleisiä tekoälybotit ovat somessa?	22
4. Paljon puhetta tekoälystä	23
Villit visiot	24
Kolme taistelevaa näkökulmaa	25
Tekoäly ja Suomi	25
5. Tekoäly – mitä meidän pitäisi tietää ja osata?	26
Algoritmit, tekoäly ja generatiivinen tekoäly	27
Mitä kansalaisen pitäisi tietää tekoälyjärjestelmistä?	27
Mitä tekoäly on?	29
Vuorovaikutus tekoälyjärjestelmien kanssa	30
Tekoälyjärjestelmien ja sovellusten käyttö	30
Tekoäly ja tietosuoja	31
Tekoällyn sudenkuopat ja eettiset kysymykset	32
Tekoäly ja etiikka	33
Ihmisen toimijuutta ja päätösvaltaa koskevat asenteet	34
6. Nuoret, oppiminen, sosiaalinen media & tekoäly	36
Lisäoppia tekoälystä	39
7. Tekoällyn ongelmat	40
Tekijänoikeudet	40
INPUT – Generatiivisen tekoällyn kouluttamiseen liittyvät kysymykset	40
OUTPUT – Generatiivisen tekoällyn tuotokset	41
Tekijänoikeudet kouluissa	41

Lähdekritiikki	41
Tietoturva	41
Läpinäkyvyyden puute	42
Virheellisen informaation leviäminen	42
Reaalimaailman ymmärtämättömyys	42
Vähemmistöjen marginalisointi	43
Ikärajasuositukset	43
Ympäristövaikutukset	43
Laadukkaat väärennökset	43
7. Tekoäly haastaa luku- ja kirjoitustaidon	45
Tekoäly voi heikentää lasten ja nuorten lukutaitoa entisestään	45
ChatGPT ja Kalle Päätalon vaimot	45
Tekeekö generatiivinen tekoäly meistä kaikista huijareita?	47
Lukutaidon merkitys korostuu tekoälyn aikakautena	48
Miten sovittaa yhteen tekoälyn sekä luku- ja kirjoitustaitojen opetus?	50
8. Mitä tarkempi syöte, sen parempi tulos	51
9. Tekoälykästä oppimista?	53
Tekoälylukutaito	54
Tekoäly on tullut jäädäkseen opetuksen avuksi	55
Tekoälyn käyttöönotto	56
Luotettavan tekoäly-ympäristön eettiset vaatimukset	57
Hybridioppimisen hyödyt	58
EdGPT: tekoälymalli koulutusta varten	58
Suosituksia tekoälyn vastuulliseen hyödyntämiseen	58
Minun tekoälyni	59
10. Pedagoginen ajattelu ja tekoälyn käyttö alakoulussa	60
Tekoäly oppimisen apuna	61
Tekoäly perustuu oppimiseen	62
11. Tekoälysovellukset lukiossa	63
Mitä Napoleon ajatteli Waterloon jälkeen?	64
Filosofian tehtäviä tekoälyllä	65
Kysely lukiolaisille	66
12. Pysäyttääkö tekoäly osaamisemme kehityksen?	68
Digitaalinen vesileima käyttöön	68
LIITE 1	70
Unescon ohjeistus koulutuksen järjestäjille	70
LIITE 2	71
Tekoälyn käytön keskeiset eettiset kysymykset opettajille	71

Sanasto

Termi	Kuvaus
Algoritmi	Prosessi tai sääntöjen kokonaisuus, jota esimerkiksi tietokone noudattaa laskelmissa tai muussa ongelmanratkaisussa. ¹
Tekoäly (AI)	<p>Tekoäly on laaja joukko tekniikoita, joissa koneet käyttävät laskennallisia kykyjä "ajatellakseen" kuten ihmiset. Tekoäly tuottaa lisäarvoa valtavasta data-aineistosta ja toimittaa sen asiakkaille pilvipalvelun kautta. Tekoälytyökaluja on monenlaisia, ja kukin niistä ratkaisee erilaisia ongelmia.²</p> <p>Teknologiaa käytetään asioiden ennustamiseen (esim. miten virus voi levitä), suositusten antamiseen (esim. mitä nettivideoita kannattaa katsoa seuraavaksi) tai päätösten tekemiseen (esim. miten essee pitäisi arvostella).</p> <p>Tekoäly voi toimia eri tavoin: noudattamalla järjestelmään syötettyjä sääntöjä, oppimalla suurista tietomääristä (esim. arvosanoista, terveydentilasta, hakuhistoriasta jne.) tai kokeilemalla ja erehtymällä.</p> <p>Tekoälyjärjestelmien käyttämät tiedot voivat olla peräisin äänitallenteista, tekstistä, kuvista tai videoista. Tekoälyjärjestelmät etsivät näistä tiedoista säännönmukaisuuksia tai toistuvia kuvioita. Kun tekoälyteknologiat löytävät kuvion, ne ovat vuorovaikutuksessa kanssamme tarjoamalla vastauksen, suosituksen tai päätöksen.³</p>
Generatiivinen tekoäly (GenAI)	Generatiivisella tekoälyllä tarkoitetaan tekoälymalleja, jotka on suunniteltu tuottamaan uutta sisältöä kirjoitetun tekstin, äänen, kuvien tai videoiden muodossa. Nykyisin monet generatiiviset tekoälyjärjestelmät hyväksyvät syötteeksi luonnollisella kielellä kirjoitettuja pyyntöjä, kehoituksia tai kysymyksiä. Esimerkiksi Chat GPT antaa kaikille verkon käyttäjille mahdollisuuden kokeilla generatiivisen tekoälyn mahdollisuuksia.
Kielimalli	<p>Kielimalleja kutsutaan usein termillä "suuri kielimalli" eli LLM (Large Language Model). Ne on koulutettu valtavalla määrällä lähinnä internetistä haettua harjoitusdataa. Kielimallin toimintaperiaate perustuu sanojen ja sanajonojen esiintymisen todennäköisyyksille. Harjoitusaineiston pohjalta se osaa ennustaa, miten käyttäjän antama tekstinpätkä jatkuu, tai tuottaa pyydetyn tyylin mukaista tekstiä. Kielimalli koostaa tietoaaineistosta tekstin, joka voi olla oikein, mutta myös vääristynyt tai virheellinen.</p> <p>Yksi esimerkki kielimallista on OpenAI:n kehittämä GPT-3, joka sisältää jopa 175 miljardia parametria.⁴</p> <p>GPT ei ole ainoa kielimalli. Muita kielimalleja ovat esimerkiksi PaLM ja Claude, sekä kuvien luomiseen keskittyvä DALL-E.</p>
ChatGPT	OpenAI julkisti marraskuussa 2022 kaikille avoimen ChatGPT-tekoälychatin, jonka taustalla on GPT-kielimalli. Kun annat tehtävän ChatGPT:lle niin se on yhteydessä GPT:hen. Sitä ohjataan ja sen kanssa voi kommunikoida luonnollisella kielellä. Tämä erottaa sen aiemmista tekoälyistä, joiden kanssa vuorovaikuttaminen on käytännössä vaatinut koodaustaitoa. ⁵
Bing Chat	Bing Chat ja moni muu ohjelmisto käyttää samaa OpenAI:n GPT-kielimallia.
Google Bard	Googlen keskusteleva tekoälysovellus Bard käyttää PaLM2-kielimallia. PaLM2 on parantunut monikielellisissä, päättelyyn liittyvissä ja koodaamiseen liittyvissä tehtävissä verrattuna edeltäjäänsä, LaMDA:aan

Ennustava tekoäly

Ennustava tekoäly on tekoälyn ala, joka keskittyy ennusteiden tekemiseen datasta. Ennustavaa tekoälyä käytetään usein päätöksenteon tukena, ja sitä voidaan käyttää monissa eri sovelluksissa, kuten esimerkiksi markkinoinnissa kohdentamaan mainontaa ja ennustamaan asiakkaiden käyttäytymistä ja terveydenhuollossa sairauksien diagnosointiin ja ennustamiseen. Ennustava tekoäly perustuu yleensä siihen, että dataa käytetään mallin luomiseen, joka voi sitten käyttää tätä mallia ennusteiden tekemiseen. Mallin luominen voi tapahtua esimerkiksi regressioanalyysin tai koneoppimisen avulla.

Syöte

Syötteellä tarkoitetaan käyttäjän tekoälylle muotoilemaa komentoa, joka voi koostua esim. tekstistä, puheesta, kuvasta tai koodista. Syötteessä olisi hyvä kertoa, mikä rooli tekoälyllä on, mitä haluaa sen tekevän ja antaa selkeät ohjeet siitä, miten edetä sekä tarpeeksi taustatietoa tehtävän suorittamiseksi. Tekoälyn syötteenä ei kannata käyttää henkilökohtaista tai salassa pidettävää aineistoa. Tekoäly hyödyntää syötteitä oppimisessa, joten kaikkea sille antamaasi tekstiä käytetään tähän tarkoitukseen ja se saattaa siten päätyä ulkopuolisten käyttöön.

- ¹ European Commission, Directorate-General for Education, Youth, Sport and Culture (2022), Final report of the Commission expert group on artificial intelligence and data in education and training, Publications Office of the European Union, <https://data.europa.eu/doi/10.2766/65087>
- ² Zhou, H. et al (2023). Accessibility Powered by AI: How artificial intelligence can help universalize access to digital content. The Scholarly kitchen. <https://scholarlykitchen.sspnet.org/2023/06/05/guest-post-accessibility-powered-by-ai-how-artificial-intelligence-can-help-universalize-access-to-digital-content/>
- ³ UNICEF (2021) AI for teens. <https://www.unicef.org/globalinsight/media/2341/file>
- ⁴ Kalmi, R. (2021). Tekoälyn uudet kielimallit hämmästyttävät. Uutishuone. <https://www.uwasa.fi/fi/uutishuone/uutiset/tekoalyn-uedet-kielimallit-hammastyttavat>
- ⁵ Katso esim. Laaksonen, S. (2023). Tekstiä tuottava tekoäly ja tieteen avoimuus. Helsingin yliopisto. <https://blogs.helsinki.fi/thinkopen/tekoaly-ja-tieteen-avoimuus/>

Johdanto

KARI KIVINEN

Tekoälyopas on jatkoa Faktabaarin vuonna 2022 julkaisemalle Digitaaliseen informaatiolukutaito-oppaalle. Julkaisun tavoitteena on taustoittaa ilmiöksi nousutta generatiivista tekoälyä eri näkökulmista.

Opas on pyritty kirjoittamaan mahdollisimman selkeästi, jotta sitä olisi kenen tahansa helppo lukea ja ymmärtää. Olemme käyttäneet runsaasti viitteitä, joiden takaa löytyy täydentävää tietoa.

Olemme tietoisia siitä, että generatiiviset tekoälypalvelut kehittyvät tavattoman nopeasti. Siksi emme ole esitelleet itse palveluita tai tuhansia sovelluksia, vaan olemme pyrkineet valitsemaan näkökulmia, jotka keskäisivät paremmin aikaa. Seuraavassa lyhyet kuvaukset julkaisun artikkeleista.

Tekoäly & faktantarkistus – Pipsa Havula

Artikkelissa esitellään ChatGPT ja muita tekoälyavusteisia ohjelmistoja, jotka voivat suorittaa erilaisia tehtäviä, kuten kirjoittaa runoja, esseitä, koodia ja tiivistelmiä sekä vastata kysymyksiin. Tekoälyohjelmat voivat myös tuottaa väärää tai harhaanjohtavaa tietoa. Lukijoita va-

roitetaan olemaan kriittisiä ja varovaisia tekoälyä käytettäessä ja tarkistamaan tekoälyn tuotosten oikeellisuus muista lähteistä. Tekoäly ei ole vielä luotettava faktantarkistaja, eivätkä tekoälyn tuotosten tunnistusohjelmat ole riittävän tehokkaita.

Tekoälyn avulla voidaan helpottaa eri tavoin virheellisen ja vääristellyn informaation tuottamista ja levittämistä. Artikkelissa esitellään muun muassa syväväärennosten käyttöä ja annetaan vinkkejä niiden tunnistamiseen.

Keskeinen ongelma on, että tekoälyä kehittävät tahot eivät kanna vastuuta virheellisistä tuotoksista eivätkä ohjelmien toimintaperiaatteet ole avoimia. Faktojen tarkistaminen jää käyttäjien tehtäväksi.

Tekoälyohjatut sosiaalisen median botit informaatiovaikuttamisessa – Harto Pönkä

Trolli- ja bottitunnuksia käytetään monenlaisen vaikuttamiseen: markkinointiin, yhteiskunnalliseen lobbaukseen, poliittiseen vaikuttamiseen tai informaatiovaikuttamiseen ja disinformaation levittämiseen. Trollien ja

bottien takana toimivia tahoja ei tavallisesti pystytä nimeämään, mutta niiden toiminnasta voidaan usein päätellä, mihin tavoitteisiin ne pyrkivät. Harto Pönkä avaa artikkelissa, miten trolleja ja botteja voidaan tunnistaa ja miten generatiivinen tekoäly muuttaa sosiaalisen median botteja.

Demokratia, osallisuus ja tekoäly - Tiina Härkönen

Demokratioiden keskiössä ovat valistuneet, mielipiteitä muodostavat yksilöt, jotka kykenevät erilaisten yhteisöjen ja yhteiskunnan jäsenenä yhteistoimintaan muiden kanssa. Ääriaineisten ja autoritaaristen valtioiden on entistä helpompi hämmentää, ohjailla keskustelua, saastuttaa informaatiota, luoda epäsopua ja jakaa ihmisiä generatiivisen tekoälyn avulla.

Euroopan unionissa painotetaan yksilöiden oikeuksia teknologioiden sääntelyssä.

Edellytykset moniääniselle osallistumiselle tekoälyn kehitykseen ovat olemassa niin lain kuin julkishallinnon eettisen ohjeistuksenkin pohjalta. Myös teknologisia ratkaisuja, joilla suuria ihmismääriä voidaan saada keskustelemaan rakentavasti yhteisistä asioista, on jo saatavilla. Esimerkki tällaisesta kokeilusta on Sitran "Mitä mieltä, Suomi?" -kampanja, jonka avulla kartoitettiin kansalaisten mielipiteitä eri asioista yhteisymmärrykseen tähtäävällä tavalla.

Muiden hyvien työkalujen tapaan erilaisten tekoälyjen tulee tuottaa arvoa meille kaikille, ei ainoastaan tekoälyjä kehittäville suuryrityksille - arvoa, joka on yhteiskunnallisesti kestävä ja ihmisoikeuksia kunnioittava nyt ja tulevaisuudessa.

Paljon puhetta tekoälystä - Minna Aslama Horowitz

Artikkelissa käsitellään tekoälyn nopeaa kehitystä ja vaikutusta yhteiskunnan ja ihmiselämän eri osa-alueisiin Median, asiantuntijoiden ja organisaatioiden näkökulmat tekoälyyn poikkeavat toisistaan, ja näkemykset vaihtelevat optimistisesta pessimistiseen. Artikkelin herättää kysymyksiä tekoälyn ymmärtämisestä, sääntelystä ja etiikasta sekä sen vaikutuksista luovuuteen, viestintään ja kulttuuriin.

Tekoälystä vallitsee kolme erilaista päänäkemystä. Tuomiopäivän julistajat varoittavat tekoälyn eksistentiaalisesta uhasta ihmiskunnalle ja maailmalle. Uudistajat kannattavat ihmiskeskeistä, vastuullista ja eettistä tekoälyn kehittämistä ja sääntelyä. Turvallisuushakuiset korostavat tekoälyn kansallisia ja kaupallisia etuja sekä riskejä. Minna Aslama Horowitzin mukaan tekoälyn merkityksestä eri sektoreille ja toimialoille Suomessa

ja maailmalla tarvittaisiin lisää pohdintaa ja vuoropuhelua. Tekoälyä koskevalle kokonaisvaltaisemmalle ja myönteisemmälle raportoinnille olisi myös paikkansa. Artikkelin kehottaa lukijoita pohtimaan, millaista tekoälyä he haluavat ja kuka siitä hyötyy.

Mitä tarkempi syöte, sen parempi tulos - Kari Kivinen

Lyhyessä artikkelissa kerrotaan syötteiden laadun merkityksestä generatiivisen tekoälypalveluiden antamiin vastauksiin. Saatua tulos riippuu paljolti siitä, miten huolella syöte tai kysymys on laadittu. Syöte tekstien taitavasta suunnittelusta (prompt engineering) on tullut tärkeä taito generatiivisten tekoälyjärjestelmien tehokkaan käytön ja käyttöönoton kannalta. Syötteessä olisi hyvä kertoa, mikä rooli tekoälyllä on ja mitä haluaa sen tekevän. Samalla tulisi antaa selkeät ohjeet siitä, miten edetä ja tarpeeksi taustatietoa tehtävän suorittamiseksi.

Tekoäly - mitä meidän pitäisi tietää ja osata? - Kari Kivinen

Artikkelissa esitellään kansalaisten digitaalisen osaamisen viitekehysten (DigComp 2.2) tekoälyliitteen osaamiskuvauksia. Mitä kansalaisten tulisi tietää tekoälystä, ja miten he voisivat hyödyntää sen mahdollisuuksia turvallisesti? Miten toimia luottavaisesti vuorovaikutuksessa tekoälyjärjestelmien ja sovellusten kanssa?

Tekoäly on jo osa kaikkien meidän arkeamme monella tapaa. Kari Kivinen pohtii tekoälyn käytön hyötyjä ja riskejä, mahdollisuuksia ja haasteita. Tekoälyn tuottamiin vastauksiin, suosituksiin ja tuloksiin kannattaa suhtautua kriittisesti, ja käyttäjän pitäisi säilyttää päätösvalta siitä, noudattaako niiden antamia suosituksia vai ei.

Nuoret, oppiminen, sosiaalinen media & tekoäly - Kari Kivinen

Artikkelissa esitellään tutkimustuloksia nuorten lisääntyneestä somen käytöstä. Siksi jokainen nuori tarvitsee kriittistä lukutaitoa eli taitoa hakea tietoa eri medioista, arvioida tekstien luotettavuutta ja vertailla tekstien sisältämää informaatiota. Sen lisäksi nuoret tarvitsevat tietoa tekoälyn haasteista ja potentiaalista ja sen turvallisesta käytöstä. Tekoälyn mahdollisuuksien tunteminen on osa yleissivistystä ja ammatillista osaamista, ja sen merkitys vain kasvaa. Artikkelin tarjoaa ehdotuksia siitä, miten hyödyntää tekoälyä opiskelussa ja miten huomioida tekoälyn käyttöön liittyvät riskit.

Tekoälyn ongelmat – Kari Kivinen

Tekoälyn haasteet-artikkelissa esitellään generatiivisten tekoälypalveluiden haasteita. Niitä ovat mm.

Tietoturva: palvelut eivät noudata EU:n tietosuojasetuksia

Läpinäkymättömyys: palvelut on koulutettu valtavalla määrällä materiaalia, joka sisältää disinformaatiota. Lisäksi palvelut hallusinoivat, eli keksivät itse asioita, joita ei ole olemassa

Tekijänoikeuskysymykset: sekä palveluiden koulutusmateriaalin keräämisessä että palveluiden tuotoksissa on lainsäädännön kannalta useita ratkaisemattomia kysymyksiä.

Reaalimaailman ymmärtämättömyys: tekoälyjärjestelmien tekstituotanto pohjautuu laskennallisiin todennäköisyyksiin, eikä niillä ole ymmärrystä generoiduista tekstistä.

Vähemmistöjen marginalisointi ja stereotyyppiset sukupuoliroolit

Ikärajasuositukset: useimmat palvelut suosittelevat 13 vuoden ikärajaa

Ympäristövaikutukset: tekoälyn kouluttaminen on resurssi-intensiivinen prosessi datan ja laskentatehon osalta.

Laadukkaat vääreennökset: syvävääreennetyt videot ovat tästä hyvä esimerkki.

Tekoäly haastaa luku- ja kirjoitustaidon – Harto Pönkä

Harto Pönkä pohtii tekoälyn vaikutuksia luku- ja kirjoitustaitoon. Somepalvelujen käytön lisääntyminen ja lukutaidon heikentyminen muodostavat viheliäisen ja vaikeasti ratkaistavan ongelman: jos käytämme yhä enemmän aikaamme sosiaalisen median sisältöjen parissa, luku- ja tiedonhakutaitomme heikkenevät, vaikka juuri somessa tarvitsimme timanttisen terävää lukutaitoa. Hän esittelee keinoja, jotka auttavat tekoälyn tuotosten arvioinnissa. Monilukutaidon hallinta näkyy kykynä toimia taitavasti ja vastuullisesti erilaisissa ympäristöissä – myös tekoälysovellusten aikana.

Tekoälykstä oppimista – Kari Kivinen

Tekoälylukutaito on varsin uusi termi, joka on vielä tarjottu kuluneen vuoden aikana generatiivisen tekoälyn räjähdysmäisen kasvun myötä.

Artikkelissa esitellään Unescon ja Euroopan komission asiantuntijaryhmän suosituksia generatiivisen tekoälypalvelujen eettisestä ja turvallisesta käytöstä koulutuksessa ja tutkimuksessa ja annetaan suosituksia sen hyödyntämisestä oppimisen tukena.

Artikkelissa kerrotaan myös uusimmista tekoälyn käyttömahdollisuuksista oppimisen tukena. Tekoälyä

voi käyttää tutorina tai valmentajana oppimisprosessin säätelyssä, jolloin oppijan ja teknologian välillä syntyy todellinen hybridisuhde.

Artikkelin liitteenä ovat Unescon ohjeistus koulutuksen järjestäjille ja Euroopan komission asiantuntijaryhmän laatimat tekoälyn käytön keskeiset eettiset kysymykset opettajille.

Pedagoginen tekoälyn käyttö alakoulussa – Saarinen

Tapani Saarinen ehdottaa, että generatiivisen tekoälyn käyttöä alakoulussa ohjaisi pedagoginen ajattelu, joka voi hyödyntää erilaisia lähtökohtia ja mahdollisuuksia. Hän esittelee neljä käytännön esimerkkiä: oppimissuunnitelmien teko moniälykkyyden teoriaa hyödyntäen, ajattelun taitojen kehittäminen argumentointia harjoittaen, tekoälyn käyttö tutorina sekä tekstin muokkaaminen oppijoiden ikäkauteen sopivaksi.

Tekoälysovellukset lukiossa – Aki Saariaho

Lukion filosofian opettaja Aki Saariaho esittelee artikkelissaan muutamia tekoälysovellusten käyttömahdollisuuksia opetuksen suunnitteluksi ja elävöittämiseksi. 75 lukiolaista vastasi hänen laatimaansa tekoälykyselyyn. Heistä 96% ilmoitti käyttävänsä tekoälypalveluja. Ainoastaan 12 % heistä luottaisi tekoälyn tekemisiin päätöksiin. Saariaho mukaan osa opiskelijoista tietää hyvin jo nyt, missä tekoälyä on ja miten se toimii, osa ei vielä hahmota, missä kaikkialla tekoälyä on jo integroitu elämäämme.

Vaarantaako tekoälyn kehittyminen ihmisen osaamisen tulevaisuuden? – Matt Cronin

Matt Cronin pohtii blogissaan tekoälypalveluiden lisääntyvän käytön seurauksia ja riippuvuutta tekoälystä. Onko riskinä, että meistä tulee oman luomuksemme orja ja emme kykene ajattelemaan tai toimimaan järkevästi ilman sen apua?

Minkä tahansa taidon oppiminen tai aiheen syvällinen ymmärtäminen vaatii paljon aikaa ja vaivaa lahjakkaaltakin oppijalta. Vaikka oppimisen prosessi on välillä stressaava, niin se on loppujen lopuksi palkitsevaa. Ensimmäistä kertaa ihmiskunnan historiassa kokonainen sukupolvi voi ohittaa tämän prosessin tekoälyn avulla ja silti edetä koulussa ja työssä. Kehitämmekö silloin tekoälyn neuroverkkoa, vai omaamme.

TM testaa – Tekoälypalvelut

Tekniikan Maailman 70-vuotisjuhlanumerossa 20/2023 julkaistiin mielenkiintoinen tekoälypalveluverailu. Verailuun osallistuivat ChatGPT 4.0, Microsoft Bing, Google Bard ja ChatGPT 3.5. Palveluista kolme jälkimmäistä ovat ilmaispalveluja.

Noin yleisesti todettiin, että "Googlen, Microsoftin ja OpenAI:n tekoälypalvelut ovat ensi vilkaisulla toistensa kopioita. Tarkempi vertailu paljastaa niistä isoja eroja. Yhteistä kaikille kuitenkin on, että ne sekoilevat välillä todella villisti".

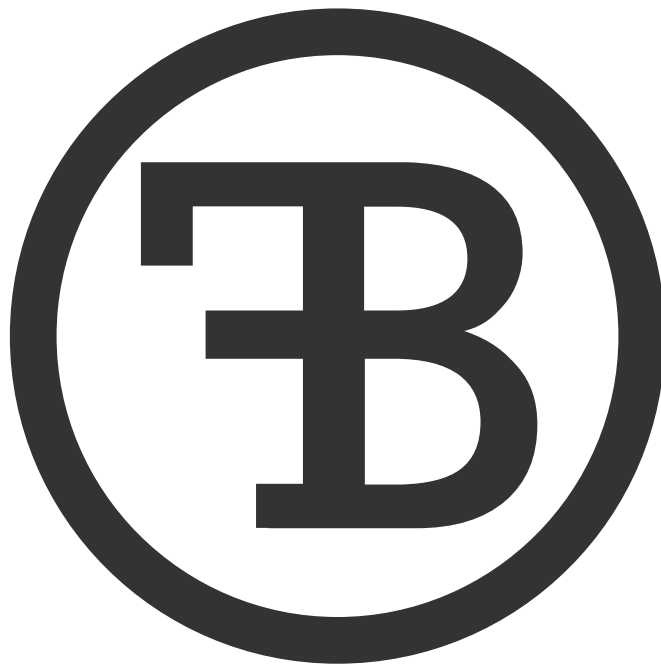
TM:n toimittaja myöntää, että tekoälyjen vertailussa on omat hankaluutensa. Ensinnäkin palvelut päivittyvät jatkuvasti. Toiseksi tekoälypalvelut pohjautuvat kielimalleihin, "joihin on sisäänrakennettuna tietynlaista satunnaisuutta. Jos samaa asiaa kysyy tekoälyltä monta kertaa peräkkäin, saa lähes aina jollain tavalla erilaisen vastauksen. Joskus vastaus on oikein ja joskus väärin, ja joskus kaikki vastaukset ovat oikeita, mutta eri tavalla muotoiltuja." Ja joskus kaikki vastaukset ovat väärin.

TM:n mukaan "Tekoälypalvelut tarjoavat kiehtovia mahdollisuuksia tehostaa tiedonhakua uusilla tavoilla. Niiden tuottamat vastaukset jättävät monilla alueilla paljon toivomisen varaa, mutta toisaalta ne kehittyvät nopeasti. Nyt on hyvä aika rohkeasti kokeilla eri palveluita, ja löytää niistä itselleen se sopivin."

* Rainisto, S. (2023). Tekoälypalvelut - selittäjien kuninkaat. Tekniikan maailma 20B. <https://tekniikanmaailma.fi/lehti/20b-2023/vertailu-tekoalypalvelut-selittajien-kuninkaat/>

Faktabaari suosittelee

- Tekoälypalvelut ovat koulutettu paljolti internetistä löydetystä materiaalista, joka sisältää paljon disinformaatiota. Lisäksi tekoälysovellukset saattavat tuottaa paikkansapitämätöntä informaatiota esimerkiksi ihmisistä, lähteistä ja tosiasioista. Tekoälyn tuottamaan aineistoon kannattaa suhtautua rakentavan kriittisesti ja tarkistaa kaikki tiedot ennen niiden käyttämistä tai jakamista.
- Tekoälypalveluiden avulla voidaan luoda laadukkaita väärennöksiä. Videoiden ja kuvien syväväärennökset ovat tästä hyvä esimerkki. Jos epäilet kuvaa tai videota väärennökseksi, niin on tärkeää pysähtyä, harkita ja kysyä itseltään kolme tärkeää kysymystä:
 - Kuka väitteen/kuvan/videon on esittänyt tai jakanut?
 - Mitä todistusaineistoa sen tueksi on esitetty?
 - Mitä muut verkkolähteet sanovat asiasta?
- Tekoälypalveluiden tietosuojakäytännöt eivät vastaa EU:n tietosuojalakeja. Siksi palveluihin ei kannata syöttää henkilötietoja eikä mitään muutakaan luottamuksellista informaatiota.



1. Tekoöly & faktantarkistus

PIPSA HAVULA

Marraskuussa 2022 yhdysvaltalainen tekoöly-yritys OpenAI julkaisi yleisölle avoimeen testaukseen ChatGPT-chatbotin. Se sai välittömästi valtavan kansainvälisen huomion, sillä ChatGPT pystyy varsin monipuolisesti muun muassa luomaan runoja ja esseitä, kirjoittamaan koodia ja antamaan neuvoja melkein mihin tahansa kysymyksiin. ChatGPT on virtuaaliavustaja ja keskustelubotti, jolta voi pyytää vaikkapa tiivistelmän talvisodasta tai kohteliaan englanninkielisen sähköpostin, jossa pahoitellaan sitä, että vastaaminen on viivästynyt. Se toimii laadukkaasti myös suomeksi.

ChatGPT ei kuitenkaan ole ainoa laatuaan. Useat yliopistot, start-up-yritykset ja teknologiajätit kehittävät parhaillaan yhä parempia ja tarkempia keskustelubotteja ja hakukoneavustajia. Artikkelin kirjoitushetkellä tunnettuja keskustelubotteja ovat ChatGPT:n lisäksi esimerkiksi Google Bard, Jasper Chat, Perplexity AI ja Bing AI. Myös Kiina kehittää omia keskustelubottejaan. Niistä tunnetuin lienee tällä hetkellä Baidun botti, jota on vaivannut sensuuriongelma.

Koska ala kehittyy kovaa vauhtia eikä kaikkien keskustelubottien ominaisuuksia ole tarkoituksenmukaista käydä tässä läpi, artikkeli tarjoaa yleiskatsauksen siihen, mitä keskustelubotit ovat, mihin niitä voi hyödyntää ja mitä ongelmia niihin liittyy faktantarkistuksen kannalta. Paras tapa ymmärtää, mistä on kyse, on käydä itse kokeilemassa jotain lukuisista keskusteluboteista.

Miten keskustelubotit toimivat?

ChatGPT ja muut vastaavat tekoölyavusteiset keskustelubotit on rakennettu suurten kielimallien päälle. Kielimalli ennustaa todennäköisyyksien perusteella sitä, miten jonkin lauseen tulisi jatkua. Esimerkiksi ChatGPT on koulutettu syöttämällä sille valtavia määriä internetistä löytyvää tekstiä, kuten lehtijuttuja, raportteja, tutkimuksia, kirjoja ja blogitekstejä. Sen jälkeen mallin toimintaa on hienosäädetty vahvistusoppimisella. ChatGPT:tä voi käyttää usealla eri kielellä, myös suomeksi.

Perplexity AI, jota voi luonnehtia ChatGPT:n ja hakukoneen yhdistelmäksi, puolestaan etsii vastauksia käyttäjän kysymyksiin reaaliaikaisesti eri puolilta verkkoa. Toisin kuin ChatGPT, Perplexity tarjoaa itse myös lähteet väitteilleen. Käyttäjä voi esittää sille jatkokysymyksiä, jolloin Perplexity vastaa niihin aiemman kontekstin perusteella. Keskustelubotit toimivatkin siten, että niille kirjoitetaan kysymys tai syöte, jonka ne tulkitsevat ja johon ne antavat sitten koulutuksensa pohjalta vastauksen. Keskustelubotteja on alettu käyttää esimerkiksi koodaamisen ja kirjoittamisen avustajana, tiivistelmien tekijänä, ideointikaverina ja oikeinkirjoituksen tarkistajana. Joillain ammattialoilla tekoölysovelluksilla voi olla vaikutuksia jopa työvoiman kysyntään.

Tekoölyavusteisten keskustelubottien ero ”perinteisiin” hakukoneisiin on se, että hakukoneesta käyttäjän

on itse etsittävä kaiken tarjotun informaation ja lukemattomien linkkien seasta relevantein tieto. Keskustelubotit sen sijaan tarjoavat informaation valmiiksi pureskellussa muodossa – joko lähteiden kanssa tai ilman. Lisäksi, kuten yllä esitettiin, niitä voi hyödyntää myös moneen muuhun asiaan kuin pelkkään tiedonhakuun.

Tekoäly ei erota, mikä on totta ja mikä ei

Pian ChatGPT:n beta-version julkaisun jälkeen moni huolestui siitä, että ChatGPT antaa välillä uskottavan kuuloisia vastauksia, jotka ovat täyttä puppua. Se saattaa keksiä lähdeviitteitä, joita ei ole olemassa, tai väittää todeksi jotakin sellaista, mitä ei ole koskaan tapahtunut. Tällaista kutsutaan kielimalleille varsin yleiseksi ”hallusinoinniksi”. Tosin hallusinointi-termiä on myös arvosteltu harhaanjohtavaksi, sillä se antaa ymmärtää, että kielimallilla olisi jokin todellisuuspohja, josta halusinaatio poikkeaa. Näin ei kuitenkaan ole. Kielimallit eivät ymmärrä sanojen merkityksiä, vaan vain sen, mikä suhde sanoilla on toisiinsa. Lauseet voivat siis olla kieliopillisesti täysin oikein ja myös kuulostaa loogisilta, mutta ne eivät pidä paikkaansa.

Esimerkiksi kysymällä ChatGPT:ltä, kuka on Ruben Stiller, se vastaa, että Stiller on suomalainen toimittaja, juontaja ja kirjailija. Se kuitenkin väittää myös, että Stiller olisi kirjoittanut useita kirjoja, ja jopa nimeää kysyttäessä neljä Stillerin kirjoittamaa kirjaa, jotka ovat kaikki keksittyjä. Stillerin syntymävuoden ChatGPT ker-

too oikein, mutta syntymäpaikan se ilmoittaa väärin. Todellisuudessa Stiller on kotoisin Nummesta. Lisäksi ChatGPT muun muassa väittää, että Stiller olisi juontanut Ruben & Joonas -ohjelmaa yhdessä Joonas Nordmanin kanssa, vaikka ohjelmaa tehtiin Joonas Hytösen kanssa. Myös väitteen siitä, että Ruben & Joonas on ollut Stillerin tunnetuin tv-ohjelma, voi kyseenalaistaa.

Tätä kirjoittaessa ChatGPT:n ilmaisversion ”tietämys” päättyy syyskuuhun 2021. Tätä uudemmissa tapah- tumista se ei osaa kertoa. Niinpä ChatGPT ei osaa esimerkiksi kertoa siitä, että Ruben Stilleristä julkaistiin elämäkerta vuonna 2022.

ChatGPT on koulutettu valtavalla tekstimäärällä, mutta se ei pysty erottamaan totuutta valheesta. Se oppii vain yhteyksiä ja säännönmukaisuuksia tekstissä, ei sitä, mikä on totta ja mikä ei. Tämän vuoksi kriittisyys botin antamia vastauksia kohtaan on äärimmäisen tärkeää.

Open AI on kertonut, että ChatGPT:llä on myös taipumus vahvistaa käyttäjien olemassa olevia uskomuksia riippumatta niiden todenperäisyydestä. Se saattaa esimerkiksi kompata käyttäjän poliittista mielipidettä, vaikka se ei perustuisi tosiasioihin.

Suuren huomion saaneella ChatGPT:llä on valuvikansa, joita on pyritty korjaamaan sen uusissa versioissa, lisäosissa ja muiden yhtiöiden kehittämässä keskusteluboteissa. Esimerkiksi Perplexity AI ja Bing AI on suunniteltu etsimään verkosta tietoa reaaliaikaisesti: niiden pitäisi siis tietää myös se, mitä viime aikoina on tapahtunut.



Kuka on Ruben Stiller?



Ruben Stiller on suomalainen toimittaja, juontaja ja kirjailija. Hän on tunnettu erityisesti televisio-ohjelmistaan, joissa hän käsittelee ajankohtaisia aiheita, haastattelee merkittäviä henkilöitä ja keskustelee yhteiskunnallisista kysymyksistä. Ruben Stiller on syntynyt 1961 Helsingissä, Suomessa.

Stiller aloitti uransa toimittajana 1980-luvulla ja on sittemmin työskennellyt useissa eri tiedotusvälineissä, kuten sanomalehdissä, radiossa ja televisiossa. Hän on tehnyt myös monia dokumenttielokuvia ja kirjoittanut useita kirjoja.

Ruben Stillerin tunnetuin televisio-ohjelma on "Ruben & Joonas" (aiemmin "Stiller"), jota hän juonsi yhdessä Joonas Nordmanin kanssa. Ohjelma oli suosittu keskusteluohjelma, jossa Stiller ja Nordman keskustelivat ajankohtaisista aiheista ja haastattelivat erilaisia vieraita.

Ruben Stiller tunnetaan myös poliittisesta aktiivisuudestaan ja avoimesta mielipiteiden ilmaisusta. Hän on ottanut kantaa erilaisiin yhteiskunnallisiin kysymyksiin ja vaikuttanut suomalaiseen keskusteluilmaperiin.

Faktabaarin testien perusteella esimerkiksi englannin kielellä toimiva hakukoneavustaja Perplexity on kohtalaisen tarkka faktoissa. Se käyttää lähteinä laatuilehtiä, kuten New York Timesia, Le Mondea, Deutsche Wellea ja luotettavia faktantarkistusorganisaatioita. Perplexityn yhtenä rajoitteena on se, että tällä hetkellä lähteiksi käyvät vain englanninkieliset materiaalit, minkä vuoksi vaikkapa monista Suomea koskevista asioista siltä on hankalaa saada kattavaa tietoa. Se ei myöskään ole yhtä monipuolinen työkalu kuin ChatGPT.

Lisäksi Perplexity antaa monimutkaisiin ja tulkinanvaraisiin kysymyksiin usein liian yksinkertaistettuja vastauksia näennäisen luotettavalla sävyllä. Se, kuten muutkin keskustelubotit, esittää yksinkertaisena totuutena asioita, joista ihmiset eivät todellisuudessa ole päässeet yhteisymmärrykseen.

Vaikka Perplexity ja monet muut keskustelubotit ovat jo varsin tarkkoja ja monella tavalla käteviä, merkittävä kysymys onkin se, millä perusteella ja mihin lähteisiin perustuen ne väittävät sitä mitä väittävät. Kuka päättää, mihin lähteisiin monimutkaisissa kysymyksissä nojataan?

Faktavirheiden lisäksi keskusteluboteilla ja hakukoneavustajilla on ”asennevinoumia”, jotka heijastelevat sitä, millä materiaalilla ja miten ne on koulutettu. Vaikka vastaukset eivät siis varsinaisesti olisi valheellisia, ne voivat olla hyvin leimaavia ja vinoutuneita esimerkiksi rotu- ja sukupuolilyksymysten, politiikan, uskonnon tai eri ammatteja koskevien stereotyyppien osalta. Nykyiset ChatGPT-mallit on koulutettu materiaaleilla, jotka heijastelevat globaalin pohjoisen arvoja ja normeja.

Tavallisesta hakukoneesta, kuten Googlesta, käyttäjä saa eteensä valtavan määrän monimutkaisen algoritmin järjestämiä linkkejä, joista täytyy osata valikoida parhaat. Keskustelubotit säästävät oman tiedonhaun ja ajattelun vaivaa, mutta toisaalta saattavat yksinkertaistaa ja tiivistää asioita liikaa. Keskustelubotin vastauksissa monet näkökulmat ja asioihin liittyvät syvemmät pohdinnat jäivät pimentoon.

Kuka tarkistaisi faktat ja kantaisi vastuun?

Yksi ChatGPT:n kaltaisten keskustelubottien huolestuttavista piirteistä on se, että disinformaation levittäjät pystyvät niiden avulla moninkertaistamaan sanomansa ja levittämään valheita valtavan tehokkaasti. Tekoälyä voi pyytää kirjoittamaan sata erilaista versiota samasta väitteestä vain muutamassa sekunnissa. Esimerkiksi Amazonin verkkokirjakauppaan on jo ilmestynyt liuta tekoälyn tehtailemia sähköisiä kirjoja, jotka kilpailevat huomiosta ja tuloista kirjailijoiden teosten kanssa. Kau-

huskenarioissa internet täyttyy heikkolaatuisesta tekoälyn kirjoittamasta hölynpölystä, josta oikean tiedon erottaminen käy mahdottomaksi. Tällaisesta informaatioasaasteesta (engl. information pollution) on puhuttu jo ennen keskustelubotteja, mutta tekoälysovellukset pysyivät pahentamaan ongelmaa huomattavasti.

Jos totuus peitellään vaihtoehtoisilla teorioilla tapah-tuneesta, vastaanottaja alkaa turtua. Kriittinen ajattelu väsyä, jos ihminen on jatkuvasti epävarma lukemiensa, kuulemiensa tai näkemiensä asioiden todenperäisyydestä. Tällaisella voi olla merkittäviä vaikutuksia demokratiaan, jonka yhtenä perustana on tiedonsaanti.

Yhdeksi ratkaisuksi ongelmaan on ehdotettu automatisoitua faktantarkistusta. Siis sitä, että tekoäly tarkistaisi tekoälyn esittämiä väitteitä.

Faktantarkistusorganisaatiot ympäri maailmaa ovat jo ennen ChatGPT:n julkaisemista kehittäneet keinoja, joiden avulla tekoälyä voitaisiin hyödyntää faktantarkistuksen tehostamisessa. Monet innovaatioista liittyvät siihen, että tekoäly auttaisi tunnistamaan väitteitä, jotka faktantarkistajien olisi syytä tarkistaa. Toistaiseksi kaiken korjaava hopealuoti, jossa tekoäly voisi luotettavasti myös tarkistaa väitteet, puuttuu. Poynter-instituutin toukokuussa 2023 tekemän testin mukaan tekoäly on ainakin tällä hetkellä varsin epäluotettava faktantarkistaja, vaikka se osuukin välillä oikeaan.

Toisaalta ratkaisuksi yritetään myös kehittää työkaluja, jotka tunnistaisivat, milloin teksti on tekoälyn kirjoittamaa. Toistaiseksi siinäkään ei ole pitävästi onnistuttu. Siitä, miten tarpeellista ylipäätään on tietää, onko teksti tekoälyn vai ihmisen tekemää, on esitetty eriäviä mielipiteitä. Onko ainoastaan sillä merkitystä, onko teksti faktuaalisesti totta vai ei?

Luottamuksen säilymisen kannalta olisi tärkeää, että joku kantaisi vastuun tekoälyn tuotoksista. Esimerkiksi ChatGPT:n kehittänyt OpenAI-yritys on omissa käyttöehdoissaan nimenomaisesti irtisanoutunut kaikesta vastuusta ChatGPT:n tuotoksista ja todennut, että chatbotin vastaukset saattavat sisältää virheitä. Myöskään esimerkiksi Microsoft ei ole vaikuttanut ottavan toisinaan Bing AI:n valheita. Yleensä keskustelubottien kehittäjät katsovatkin, että on teknologioiden käyttäjien vastuulla varmistaa, että chatbottien käyttämisen vaikutus on pikemminkin myönteinen kuin kielteinen. Käyttäjä itse on vastuussa faktojen tarkistamisesta.

Moni asiantuntija kuitenkin näkee, että keskustelubottien kehittäjien pitäisi kantaa vastuuta toimistaan ja päätöksistään. Keskustelubotteja pitäisi kehittää avoimesti ja läpinäkyvästi, ja niissä pitäisi olla selkeät mekanismit virheiden käsittelemiseen. Teknologioiden kehittäjillä on myös toivottu läpinäkyvyyttä sen suhteen, mitä harjoitusdataa keskustelubotteihin on käytetty ja miten ne on käytännössä koulutettu.

Syväväärennökset

Oma lukunsa netin informaatioympäristön hämmentämisessä ja sotkemisessa ovat vielä syväväärennökset (deepfake). Generatiivisen tekoälyn avulla voidaan nykyisin hyvin todentuntuisesti muuttaa ihmisten kasvoja, vartaloa ja ääntä siten, että heidät saadaan tekemään ja sanomaan asioita, joita he eivät todellisuudessa ole koskaan tehneet tai sanoneet. Tekoälyllä voidaan myös helposti ja nopeasti luoda aitojen valokuvien näköisiä kuvia lähes mistä tahansa. Videoihin ja kuviin voidaan lisäksi vaikkapa lisätä esineitä totuuden vääristämiseksi.

Samoihin aikoihin, kun laajat kielimallit ovat tulleet kaikkien käytettäviksi, myös syväväärennökset ovat lisääntyneet räjähdysmäisesti. Niiden tuloa on povattu ja ennustettu jo vuosikausia, mutta nyt ne näyttävät viimein yleistyneen, kun kuvien tekemiseen tarkoitettujen tekoälysovellukset, kuten Midjourney ja DALL-E, sekä lukuisat kaupalliset äänenkloonausohjelmat on julkaistu kenen tahansa netinkäyttäjän saataville.

Laajat kielimallit osaltaan myös edesauttavat hyvien syväväärennösten luomista, sillä kuka tahansa voi tehdä nyt kieliopillisesti uskottavaa tekstiä pienemminkin kielillä, kuten suomeksi. Teksti puolestaan voidaan muuntaa puheeksi, joka vieläpä saadaan halutessa kuulostamaan vaikkapa presidentti Sauli Niinistöltä.

Jo vuosikausia etenkin naisia ympäri maailmaa on kiristetty ja häpäisty tekemällä heistä syväväärennettyä pornoa, mutta vuonna 2023 deepfake-teknologiaa on alettu ahkerasti käyttää myös poliittisessa kamppailussa ja vaaleissa länsimaissa. Esimerkiksi Slovakian parlamenttivaaleissa syksyllä 2023 generatiivista tekoälyä käytettiin useita kertoja vääristelemään poliitikkojen sanomisia. Erästä maan kärkipoliitikosta oli esimerkiksi muokattu ääninauha, jossa kuulosti siltä kuin tämä olisi keskustellut toimittajan kanssa vaalien manipuloinnista ja äänten ostamisesta Slovakian romanivähemmistöltä. Todellisuudessa ääninauha oli deepfake-teknologialla tehty.

Yhdysvalloissa Floridan kuvernööri Ron DeSantis joutui syksyllä 2023 syväväärennöksen uhriksi, kun verkossa levisi hänestä tekoälyllä tehty video, jossa DeSantis kertoi luopuvansa presidenttiehdokkuudestaan.

Helmikuussa Chicagon pormestariehdokkaiden esivaalissa joku puolestaan oli kloonannut ehdokas Paul Vallasin äänen ja laittanut hänen suuhunsa sanoja, joiden perusteella vaikutti siltä kuin Vallas hyväksyisi poliisiväkivallan.

Deepfake-teknologialla levitetään nykyisin lähes poikkeuksetta disinformaatiota suurten uutistapahtumien aikaan.

Suomikaan ei ole näistä kehityskuluista irrallaan. Täälläkin esimerkiksi Tiktokin ja Youtuben käyttäjät ovat kloonanneet kotimaisten poliitikkojen ääniä ja hassutelleet sillä, mitä kaikkea poliitikot voidaan saada sanomaan ja laulamaan. Huoltovarmuuskeskuksen johtava varautumisasiantuntija Antti Sillanpää onkin arvioinut Suomen Kuvalehdessä, että ”hyvin ajoitettu, ajatuksia sekoittava syväväärennös voi vaikuttaa äänestyspäätökseen.”

Samalla on konkretisoitunut myös toinen huoli syväväärennöksistä: se, että myös aitoa materiaalia on alettu syyttää väärennetyksi. Tällä voi olla vakavia vaikutuksia oikean tiedon leviämiseen ja siihen uskomiseen. Tästäkin on jo lukuisia esimerkkejä, tuore tapaus esimerkiksi Israelin ja Gazan konfliktista.

Muutamia päiviä Hamasin hyökkäyksen jälkeen Israelin pääministerin toimisto julkaisi X:ssä kuvia, joiden se kertoi olevan otettu Hamasin murhaamista ja polttamista israelilaisvauvoista. Kuvissa näkyi valkoisessa kangaskääreessä jonkinlaisia palaneita jätteitä. Suomenkäyttäjät olivat ladanneet näitä kuvia tekoälykuvia tunnistavaan AI or Not -sovellukseen, jonka johtopäätös oli se, että kuvat ovat luultavasti tekoälyllä tehtyjä.

Esimerkiksi Al Jazeera Arabic uutisoi asiasta väittäen, että Israel oli julkaissut tekoälyllä manipuloituja kuvia. Samaan aikaan lisähämmennystä toi se, että joku oli muokannut yhtä Israelin julkaisemista kuvista niin, että kangaskääreessä olikin nyt koiranpentu. Koiranpentukuvaa levitettiin väittäen, että se oli aito ja alkuperäinen kuva, jota Israelin viranomaiset olivat muokanneet valehdellakseen Hamasin toimista. Tapauksesta on kattavasti raportoinut France24.

Lopulta faktantarkistajien ja syväväärennösasiantuntijoiden johtopäätös oli se, että Israelin julkaisemat kuvat olivat aitoja. Moni kuitenkin jo kuvitteli, että Israelin pääministerin toimisto levittää syväväärennöksiä. Tapaus nosti konkreettisesti esiin tekoälykuvien ja -videoiden tunnistamisteknologian vaarallisen sudenkuopan: sen, että ne antavat vääriä positiivisia. Onkin tärkeää ymmärtää, että syväväärennösten vastateknologia, tunnistusteknologia, ei ole täydellistä ainakaan tällä hetkellä. Tässäkään mielessä tekoäly ei toistaiseksi ole luotettava faktantarkistaja.

Jotta syväväärennöksillä voidaan saada vahinkoa aikaan, niitä pitää pystyä levittämään viestipalveluissa ja sosiaalisessa mediassa. Digijäteiltä, kuten Metalta, Tiktokilta, Googlelta ja Microsoftilta onkin vaadittu vastuunkantoa syväväärennösten leviämisen suitsemiseksi.

Vaikka taitavasti tehtyjä syväväärennöksiä onkin liki mahdotonta erottaa aidosta materiaalista, joistain asioista niitä voi edelleen yrittää tunnistaa:

Yleispätevinä ohjeina voi sanoa, että:

- Jos jokin somessa vastaan tuleva kuva, video tai ääninauha herättää voimakkaan tunnereaktion, ensimmäiseksi on syytä pysähtyä, harkita ja harkita vielä uudestaan ennen kuin uskoo materiaalia ja jakaa sitä eteenpäin.
- Pohdi kolmea kysymystä: Kuka kuvan/videon/ääninauhahan on julkaissut? Mitä todistusaineistoa sen tueksi on esitetty? Mitä muut lähteet sanovat asiasta?
- Käytä lateraalista lukutapaa: Ovatko luotettavat lähteet raportoineet asiasta? Onko joku kyseenalaistanut materiaalin aitouden kommenteissa? Löydätkö muita kuvia/videoita/materiaalia samasta tapahtumasta?
- Ole skeptinen etenkin, jos juuri ennen vaaleja sosiaalisessa mediassa alkaa levitä jokin skandaalinkäryinen kuva, video tai ääninauha. Ota tarpeen mukaan yhteyttä viranomaisiin, faktantarkistajiin tai muuhun journalistiseen mediaan.

Tiettyjen yksityiskohtien kanssa generatiivisella tekoälyllä on edelleen usein haasteita. Sen vuoksi:

- Tarkista, ovatko videolla puhuvan ihmisen suun liikkeet ja ääni yhtäaikaiset ja uskottavat.
- Kuulostaako ääninauhalla siltä, että puhuja hengittää välissä? Tulevatko hengitykset loogisissa kohdissa? Entä onko puheessa täytesanoja ja epäröintiä, vai kuulostaako puhe epäilyttävän harkitulta ja suoraviivaiselta?
- Onko puheen intonaatio liian monotonista olakseen totta?
- Räpsyvätkö puhujan silmät? Entä katsovatko ne samaan suuntaan?
- Tekoälyllä tehdyissä kuvissa ja videoissa virheitä voi näkyä esimerkiksi käsissä, hampaissa, korvissa, silmälasissa ja vartalon mittasuhteissa. Ihmisen iho myös usein näyttää liian pehmeältä olakseen totta. Tarkista ne.

⁷ Google Bard <https://bard.google.com/>; Jasper AI <https://www.jasper.ai/chat>; Perplexity AI <https://www.perplexity.ai/>; Bing AI <https://www.bing.com/>

⁸ McDonnell, Stephen (2023). Elusive Ernie: China's new chatbot has a censorship problem. BBC. <https://www.bbc.com/news/world-asia-66727459>

⁹ Napolitano, Elizabeth (2023). AI eliminated nearly 4,000 jobs in May, report says. CBS News. <https://www.cbsnews.com/news/ai-job-losses-artificial-intelligence-challenger-report/>

¹⁰ OpenAI (luettu 19.12.2023). Is ChatGPT biased? <https://help.openai.com/en/articles/8313359-is-chatgpt-biased>

¹¹ Damisch J., Kim F. & Shi H. (2023). Reducing Bias in Large Language Models. Berkeley School of Information. <https://www.ischool.berkeley.edu/projects/2023/reducing-bias-large-language-models>

¹² Larraz Irene & Lozano Borja (2023). How AI-generated content could both fuel disinformation and improve fact-checking. Poynter. <https://www.poynter.org/commentary/2023/artificial-intelligence-help-harm-fact-checking/>

¹³ Brewster, J., Fishman, Z. & Xu, E. (2023). Funding the Next Generation of Content Farms: Some of the World's Largest Blue Chip Brands Unintentionally Support the Spread of Unreliable AI-Generated News Websites. NewsGuard. <https://www.newsguardtech.com/misinformation-monitor/june-2023/>

¹⁴ Roscoe, Jules (2023). AI-Generated Books of Nonsense Are All Over Amazon's Bestseller Lists. Vice. <https://www.vice.com/en/article/v7b774/ai-generated-books-of-nonsense-are-all-over-amazons-bestseller-lists>

¹⁵ Abels, Grace (2022). What is the future of automated fact-checking? Fact-checkers discuss. Poynter. <https://www.poynter.org/fact-checking/2022/how-will-automated-fact-checking-work/>

¹⁶ Abels, Grace (2023). Can ChatGPT fact-check? We tested. Poynter. <https://www.poynter.org/fact-checking/2023/chatgpt-ai-replace-fact-checking/>

¹⁷ Algorithm Watch (2023). ChatGPT and Co: Are AI-driven search engines a threat to democratic elections? <https://algorithmwatch.org/en/bing-chat-election-2023/>

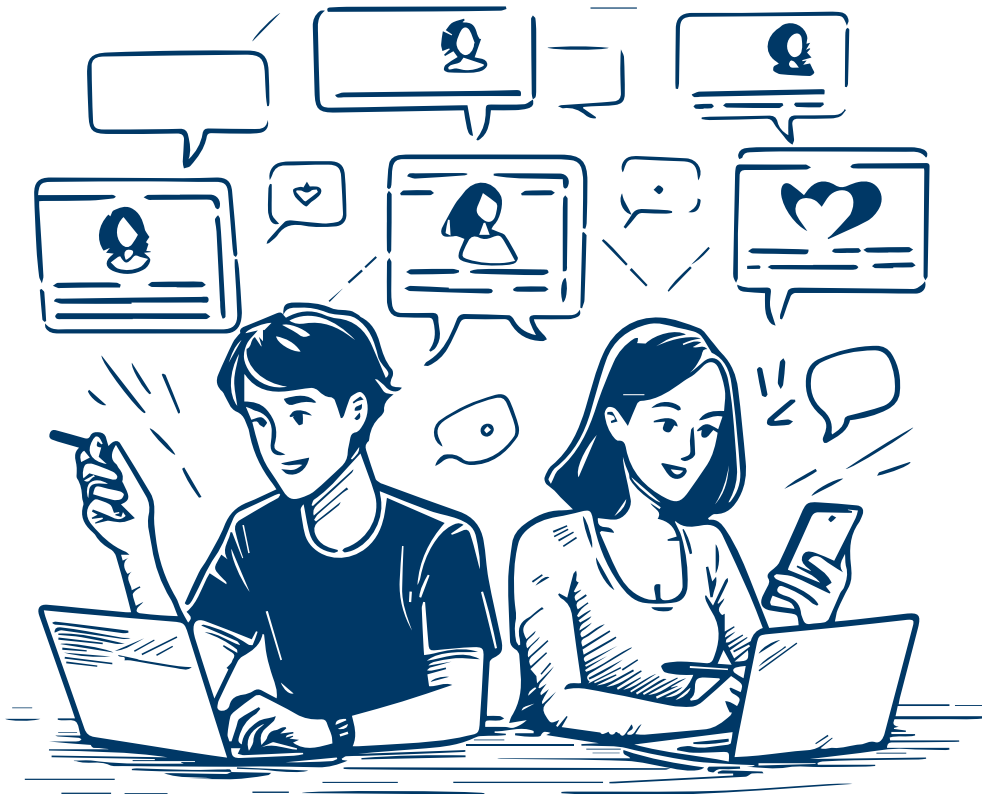
¹⁸ Meager, Morgan (2023). Slovakia's Election Deepfakes Show AI Is a Danger to Democracy Wired. <https://www.wired.co.uk/article/slovakia-election-deepfakes>

¹⁹ Suci, Peter (2023). There Is Now A Deep Fake Video Of Ron DeSantis Dropping Out Of The 2024 Race. Forbes.com. <https://www.forbes.com/sites/petersuci/2023/09/02/there-is-now-a-deep-fake-video-of-ron-desantis-dropping-out-of-the-2024-race/?sh=287ed097bf9>

²⁰ Vallas Campaign denounces fake, deceptive impersonation video (katsottu 19.12.2023). <https://www.paulvallas2023.com/post/vallas-campaign-denounces-fake-deceptive-impersonation-video>

²¹ Havula, Pipsa (2023). Hölynpölyä. Suomen Kuvalehti. <https://suomenkuvalehti.fi/uutisviikko/huijausten-kanssa-syyta-olla-tarkkana-vaalien-alla-hyvin-ajoitettu-ajatuksia-sekoittavasyvavaarennos-voi-vaikuttaa-aanestyspaatokseen/?shared=1273770-c6709861-500>

²² Gallo, Nathan (2023). Was this photo of a dead Israeli baby AI-generated? When AI-detection errors muddle public debate. France24. <https://observers.france24.com/en/middle-east/20231018-was-this-photo-of-a-dead-israeli-baby-ai-generated-when-ai-detection-errors-muddle-public-debate>



2. Demokratia, osallisuus ja tekoäly

TIINA HÄRKÖNEN, SITRA

Demokratiassa valta on kansalla – siis meillä ihan tavallisilla ihmisillä. Se tarkoittaa muun muassa oikeutta osallistua yhteiskunnalliseen päätöksentekoon joko suoraan tai valitsemalla edustajia, jotka laativat lakeja ja tekevät päätöksiä puolestamme, parhaan käsityksensä mukaan. Suomen tulee demokraattisena valtiona kunnioittaa lukuisia eri perus- ja ihmisoikeuksia, kuten sananvapautta ja tasa-arvoa.

Yhteiskunnassa tapahtuu kuitenkin valtavasti arkeamme vaikuttavia asioita, joihin meillä on joko hyvin vähän tai ei lainkaan mahdollisuutta vaikuttaa. Teknologinen kehitys kuuluu näihin: generatiivista tekoälyä on kehitetty tavallisten ihmisten ulottumattomissa, ja se on uinut lukemattomiin eri palveluihin ja käyttötarkoituksiin kuin huomaamatta, ajan saatossa. Tästä huolimatta sen kehitys olisi ollut täysin mahdotonta ilman tavallisia ihmisiä, jotka käyttävät arjessaan digitaalisia palveluita. Heiltä tulee tekoälyn opettamiseen tarvittava data, tavalla tai toisella, luvatta tai luvan kanssa.

Demokratioiden keskiössä ovat valistuneet, mielipiteitä muodostavat yksilöt, jotka kykenevät erilaisten yhteisöjen ja yhteiskunnan jäsenenä yhteistoimintaan muiden kanssa. Generatiivisen tekoälyn mukaantulo tarjoaa mahdollisuuksia haastaa tätä perustaa yhä enemmän. Ääriaineisten ja autoritaaristen valtioiden on entistä helpompi hämmentää, ohjailla keskustelua, saastuttaa informaatiota, luoda epäsopua ja jakaa ihmisiä.

Luottamus toisiin ihmisiin ja yleisesti yhteiskuntaan on demokratioissa täysin välttämätöntä. Jos mihinkään tai kehenkään ei voi luottaa, on demokratia, ja samalla yksilönvapaus, todellisessa vaarassa. Siksi tähän kehitykseen on tartuttava määrätietoisesti, kaikilla yhteiskunnan alueilla ja tasoilla.

Koska digitalisaation tuloksena syntynyt tekoäly vaikuttaa elämäämme monin eri tavoin, on kohtuullista ajatella, että sen hyödyntäjillä ja käyttäjillä – tai sen ongelmista kärsivillä – tulisi puolestaan olla mahdollisuus vaikuttaa tekoälyn käyttöön ja kehitykseen. Uudet teknologiat kulkevat käsi kädessä demokratian ja yksilönvapauksien kanssa.

EU:n lait suojelevat ihmisiä ja organisaatioita

Euroopan unionissa painotetaan yksilöiden oikeuksia teknologioiden sääntelyssä. Tulevan tekoölysäätelyn (AI Act) tavoitteena on varmistaa, että EU:ssa käytettävät tekoölyjärjestelmät ovat turvallisia, läpinäkyviä, jäljitettäviä, syrjimättömiä ja ympäristöystävällisiä. Näkemys on myös, että tekoölyjärjestelmiä tulisi valvoa ihmisten eikä automaation avulla, jotta haitallisia vaikutuksia voitaisiin paremmin hallita.

EU:n dataan liittyvä sääntely turvaa yksilöiden ja erityisesti lasten oikeuksia jo melko hyvin. Myös uusi digipalvelusäädös (Digital Services Act, DSA) jatkaa joulukuusta eurooppalaisen tietosuojasäädöksen (GDPR) linjoilla. DSA tuo kuitenkin mukanaan lisää velvoitteita suurille palveluntarjoajille asettaessaan muun muassa täyskiellon lasten profiloinnille mainontaa varten.

Jo GDPR:n voimaantulo edellytti, että informaatio siitä miten ja miksi dataa kerätään, pitää tarjota myös lapsille ymmärrettävässä muodossa. Nyt se on sitäkin tärkeämpää, kun dataa yksilöistä ei enää kerätä vain ”mainosten kohdistamista varten”, vaan sitä käytetään tekoölyn kehittämiseen.

Vaikka tavoitteena onkin esimerkiksi huolehtia eurooppalaisesta tekoölyosaamisesta ja yritysten innovaatiokyvystä, eurooppalaisella tekoölylainsäädännöllä halutaan myös varmistaa, että käytettävät teknologiat ovat ihmiskeskeisiä ja luotettavia.

Matti Meikäläisen kannalta tämä tarkoittaa muun muassa sitä, että alustoja on kielletty manipuloimasta ja taivuttelemasta käyttäjiään niin, että heidän kykynsä tehdä valistuneita ja vapaita valintoja estyy²². Heitä ei saa myöskään pisteyttää sosiaalisesti, millä tarkoitetaan ihmisten luokittelua käyttäytymisen, taloudellisen aseman tai henkilökohtaisten ominaisuuksien perusteella. Sosiaalista pisteytystä käytetään jo Kiinassa esimerkiksi päätettäessä siitä, annetaanko kansalaisille lupaa matkustaa ulkomaille.

Suomi, päättäjät ja tekoöly

Valtiovarainministeriö sanoo tekoölyn eettisessä ohjeistuksessaan näin: ”Julkisessa hallinnossa on kyse toimimisesta kansalaisten valtuutuksella, julkisilla varoilla ja kansalaisten eduksi. Tämä laillinen ja eettinen suhde kansalaisen, valtion ja virkahenkilön välillä on myös tekoölyn käytön perusta.”²³

On tärkeää huomata, että kansalaisten valtuutuksella toimiminen voi tarkoittaa hyvin erilaisia asioita, eikä Suomessa ole toistaiseksi käytäntöä ottaa kansalaisia mukaan aktiivisesti keskustelemaan ja kehittämään tekoölyä julkisten palveluiden yhteydessä.

Edellytykset moniääniselle osallistumiselle tekoölyn kehitykseen ovat olemassa niin lain kuin julkishallinnon eettisen ohjeistuksenkin pohjalta. Myös teknolo-

gisia ratkaisuja, joilla suuria ihmismääriä voidaan saada keskustelemaan rakentavasti yhteisistä asioista, on jo saatavilla. Esimerkki tällaisesta kokeilusta on Sitran ”Mitä mieltä, Suomi?” -kampanja, jonka avulla kartoitettiin kansalaisten mielipiteitä eri asioista yhteisymmärrykseen tähtävällä tavalla.²⁴

Digitalisaation, mutta myös tekoölyn hyödyntäminen sosiaali- ja terveystaloudessa on välttämätöntä, jotta palvelut voidaan taata kaikille. Sosiaali- ja terveysministeriö onkin näyttänyt esimerkkiä eri ryhmien osallistamisesta suunnitteluun ja kysynyt myös kansalaisten näkemyksiä siitä, millaisia digitaalisia palveluja he haluavat tulevaisuudessa.²⁵ Näin toimimalla voitaisiin varmistaa, että myös tekoölyteknologioilla edistetään hallinnon yleisiä tavoitteita, kuten avointa ja aktiivista kansalaisuutta tukevaa hallintoa ja luottamukseen ja perusoikeuksien edistämiseen perustuvaa yhteiskuntaa.

Kuka saa vaikuttaa tekoölyn kehitykseen?

Jos sinä saisit vaikuttaa tekoölyn tulevaisuuteen, mitä siltä toivoisit? Kysymys sinulle, olitpa kuka vain, on yhtä tärkeä kuin kysymys tekoölytutkijalle tai -kehittäjälle. Juuri monet tekoölytutkijat ovat viime aikoina toivoneet enemmän moniäänisyyttä tekoölykeskusteluun, sillä olemme painottaneet liikaa teknologiaa, jolloin ihminen on ehkä jäänyt taka-alalle.

Tällä hetkellä toiveenasi voisi olla niinkin yksinkertainen asia kuin että käyttämäsi generatiivinen tekoöly ei hallusinoisi vaan sen antamaan tietoon voisi luottaa. Jos katsomme vähän pidemmälle tulevaisuuteen, kysymys on paljon suuremmasta ja tärkeämmästä asiasta. On tärkeää pohtia, kuka käyttää valtaa meihin ja ympäristöömme, mutta myös sitä, onko meillä itsellämme lainkaan valtaa vaikuttaa tulevaisuuden suuntaan?

Tekoölyjä on monenlaisia, joten niin on luonnollisesti tapoja hyödyntääkin niitä. Siksi vaikutus yhteiskuntaan tai yksilöiden elämään on joissakin tapauksissa näkyvämpää ja jää toisaalla täysin pimentoon. Tekoöly voi olla vaikkapa tavallisen ihmisen viihdettä tai tiedonhaku keskustellen chatbotin välityksellä. Toisaalla se on yrityksen työkalu, jonka avulla voi summata tietoa tai kääntää asiantuntija-artikkeleita. Tai sitten se on täysin näkymättömissä, konepellin alla, esimerkiksi vakuutusyhtiön tai viraston koneistossa, verkkokaupan tai somen suosituksissa, jossa se tekee ”itsenäisiä” päätöksiä henkilökohtaisiin tietoihin pohjautuen. Silloin tekoölyyn ei voi suoraan itse vaikuttaa, vaikka sillä on suuri merkitys ihmisen arjessa.

Digitalisaatio on tuonut monien hyötyjen ohella paljon haittoja. Yksi tärkeimmistä ongelmista on digitaalisten varantojen (infrastrukturi, data ja osaaminen) keskittyminen pienelle joukolla maailmanlaajuisia jättiyrityksiä. Generatiivisen tekoölyn kehityksessä on hyödynnetty suurten ihmisjoukkojen käyttämistä digi-

taalisista palveluista – sosiaalisesta mediasta, verkko-kaupoista ja hakukoneista – kerättyä dataa, joka on digitaalisille jäteille arvokasta raaka-ainetta. Siksi suurimmat voittajat ovat ainakin toist likimain samoja yrityksiä kuin aiemminkin.

Jos emme aktiivisesti toimi eri puolilla yhteiskuntaa, äärimmäisessä tapauksessa muutama yksilö suuryritysten johdossa voisi saada sellaista valtaa, jollaista ei ole koskaan ennen nähty ihmiskunnan historiassa.²⁶ Se olisi räikeässä ristiriidassa demokratian periaatteiden kanssa. Jos digitaalinen valta keskittyy edelleen, keskittyy valitettavasti myös valta vaikuttaa tulevaisuuteen, ja se olisi kaikkien kannalta huono asia.

Lait ja oman ymmärryksemme kasvattaminen suojelevat meitä jonkin verran haitoilta, mutta entä hyödyt: kuinka ottaa kaikki irti lupaavasta teknologiasta, jolla voisimme oikeasti parantaa maailmaa? Tekoälyn kehittämisen valta on saatava hajautettua ja vieläpä meille yksilöille asti.

Joissakin maissa ratkaisua on haettu hallinnon tekemän tekoälystrategian ja ihmisten toimijuuden yhdistelmällä.

Mainio esimerkki löytyy Skotlannista, jossa maan hallinto on linjannut tekoälystrategiansa olevan luottamuksen arvoinen, eettinen ja osallistava.²⁷ Sana ”osallistava” ei ole jäänyt vain kirjoitettuun dokumenttiin, vaan sitä toteutetaan määrätietoisella otteella. Skotlannissa tekoälyjärjestelmät valjastetaan demokratian palvelukseen, sillä yksi kehittämisen periaatteista edellyttää, että tekoälyjärjestelmät suunnitellaan kunnioittamaan oikeusvaltioperiaatetta, ihmisoikeuksia, demokraattisia arvoja ja monimuotoisuutta.

Ilahduttavinta on se, että myös lapset ja nuoret pääsevät osallistumaan, toimimaan ja siten vaikuttamaan. Huolimatta tekoälyn nopeasta kehityksestä, lasten kokemuksia tekoälystä on tutkittu vain vähän. Skotlantilaisten kokeilukoulujen tutkimusohjelmassa lapset pohtivat tekoälyn eettisiä kysymyksiä. Ohjelma kannustaa lapsia toimijuuteen. Lasten huolet, toiveet ja näkemykset tekoälystä koetaan arvokkaiksi ja ne viedään päättäjille pohdittaviksi tekoälyä koskevan päätöksen pohjaksi. Lapsen oikeudet ovat olennainen osa tekoälyajattelua, jolloin lapsille ja nuorille annetaan tulevaisuusvaltaa vaikuttaa itseään koskeviin asioihin.

Haaveena demokraattinen ja inhimillisiin arvoihin perustuva tekoäly

Meillä kaikilla on teoriassa oikeus vaikuttaa tulevaisuuteen, mutta kuinka käytännössä osallistua ”kaukana olevien” asioiden, kuten tekoälyinnovaatioiden, -po-

litiikan ja -hallinnon muotoiluun? Aivan ensimmäiseksi jokaisen täytyy tunnistaa olevansa oikeutettu mielipiteeseen tekoälystä ilman, että se edellyttää teknistä osaamista.

Vaikka tekoälyn käyttäminen sopivien lääkkeiden kehittämiseen voi olla kaukana tavallisen ihmisen elämästä, tulisi yksilöiden saada vaikuttaa kaikkeen teknologian kehitystyöhön, jolla vaikutetaan ihmisten oikeuksiin ja elämään.

Tekoälyn käyttöä tarkastellaan jo nyt kouluissa sekä opetusvälineen roolissa että muutoinkin. Koulut ovat hyvin eri vaiheissa generatiivisen tekoälyn hyödyntämisessä ja opettamisessa, mutta kuka tahansa voi pohtia tekoälyn merkitystä vaikkapa yhteiskuntaopin, psykologian tai uskonnon tunneilla. Vaikka otamme vasta ensi askeleita tekoälyopetuksessa, on hyvä pohtia myös, kuinka sen mukaantulo voi muuttaa koulua ja kaikkea opetusta.

The Child Friendly Governance Project on kirjoittanut, että nykyinen yhteiskuntamme on hyvä kasvatamaan kuluttajia, mutta ei niin hyvä kasvattamaan kansalaisia. Jos pääsemme kouluissa paneutumaan tekoälyjen etiikkaan, demokratiaan ja ihmisoikeuksiin sekä sen moninaisiin yhteiskunnallisiin vaikutuksiin, varustamme lapsia ja nuoria paremmin tulevaisuuteen. Demokratia- ja ihmisoikeuskasvatuksen tulisi tukea tekoälyopetusta ja päinvastoin.

Sosiaalibiologi E. O. Wilson on puolestaan (vapaasti suomentaen) todennut, että ihmiskunnan todellinen ongelma on se, että tunne-elämämme on vielä kivikautista, instituutiomme pohjaavat keskiaikaan ja teknologia on jumalankaltaisella tasolla.²⁸ Sekä tekoälyn että muun teknologian kehittämisen kannalta on tärkeää, että kaikki kolme aluetta ovat tasapainossa.

Ongelman ratkaisu edellyttää, että rakennamme yhdessä koko yhteiskuntana yhteisymmärrystä siitä, millainen on tulevien tekoälyjen ”perustuslaki”, johon voivat sitoutua niin yksilöt, yritykset kuin julkinen valtakain. Kokeiluja maailmalta jo löytyy. Esimerkiksi Yhdysvalloissa tuhannelta tavalliselta ihmiseltä kysyttiin, millainen heidän mielestään olisi ihmisten kannalta hyvä ja toimiva tekoäly. Kansainvälisen Forus-verkoston kokeilusta syntyi kansalaisyhteiskunnan manifesti tekoälyperustuslaiksi, joka julkaistiin 6. joulukuuta 2023.²⁹

Mielenkiintoisista uusista teknologioista ja sovelluksista huolimatta vauhtisokeus ei saa hämärtää tekoälyn roolia ihmisten ja organisaatioiden työkaluna, tukena sellaisissa vaativissa asioissa, joissa teknologialla on paikkansa. Muiden hyvien työkalujen tapaan erilaisten tekoälyjen tulee tuottaa arvoa meille kaikille, ei ainoastaan tekoälyjä kehittäville suuryrityksille – arvoa, joka on yhteiskunnallisesti kestävä ja ihmisoikeuksia kunnioittava nyt ja tulevaisuudessa.

²³ European Commission (2023) Questions and answers: Digital Services Act. https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_2348

²⁴ Valtionvarainministeriö (ladattu 19.12.2023) Tekoälyn eettinen ohjeistus. <https://vm.fi/tekoalyn-eettinen-ohjeistus>

²⁵ Sitra (ladattu 19.12.2023) Mikä ihmeen ”Mitä mieltä, Suomi?” <https://www.sitra.fi/hankkeet/mitamieltasuomi/#mista-on-kyse>

²⁶ Valtioneuvosto (ladattu 19.12.2023) Sosiaali- ja terveydenhuollon digitalisaation ja tiedonhallinnan strategia painottaa joustavuutta. <https://valtioneuvosto.fi/-/1271139/>

sosiaali- ja terveydenhuollon-digitalisaation- ja-tiedonhallinnan-strategia-painottaa-joustavuutta

²⁷ Scottish AI Alliance (2021). Scotland's AI strategy. <https://www.scotlandaisstrategy.com/the-strategy>

²⁸ Oxford reference. <https://www.oxfordreference.com/display/10.1093/acref/9780191826719.001.0001/q-oro-ed4-00016553>

²⁹ Forus (2023) Civil society manifesto for ethical AI. <https://www.forus-international.org/en/event-detail/9094-civil-society-manifesto-for-ethical-ai-launch-event>



3. Tekoälyohjatut sosiaalisen median botit informaatiovaikuttamisessa

HARTO PÖNKÄ

Trolli- ja bottitunnukset ovat monen sosiaalisen median palvelun vitsaus. Yhteistä niille on epäaito toiminta: ne saattavat esimerkiksi manipuloida tiettyjä aiheita julkaisemalla runsaasti niihin liittyvää sisältöä tai kommentteja. Usein botti- ja trollitunnukset on luotu siten, että ne näyttävät tavallisilta käyttäjiltä. Esimerkiksi profiilikuvat on voitu generoida tekoälyn avulla ja nimetä tunnukset yleisillä etu- ja sukunimillä.

Trolli- ja bottitunnuksia voidaan käyttää monenlaisen vaikuttamiseen: markkinointiin, yhteiskunnalliseen lobbaamiseen, poliittiseen vaikuttamiseen tai informaatiovaikuttamiseen ja disinformaation levittämiseen. Trollien ja bottien takana toimivia tahoja ei tavallisesti

pystytä nimeämään, mutta niiden toiminnasta voidaan usein päätellä, mihin tavoitteisiin ne pyrkivät.

Mitä eroa on trollilla ja botilla?

Sanalla "trolli" on useita merkityksiä. Jo ennen nykyisiä sosiaalisen median palveluita trolleilla viitattiin netin keskustelufoorumeilla käyttäjiin, jotka pyrkivät ärsyttämään muita. Tuolloin syntyi edelleen hyvä neuvo olla "ruokkimatta trollia", millä varoitetaan siitä, että jotkut keskustelijat saavat itselleen hupia provosoimalla muita loputtomilta tuntuviin väittelyihin kanssaan. Tätä voi edelleen pitää trolloamisen yleisimpänä merkityksenä.

Erityisesti 2010-luvulta lähtien trolli-sana on saanut huomattavasti aiempaa vakavampia merkityksiä. Tänä päivänä trolleilla viitataan yleisesti tunnuksiin, jotka leviävät tietyn aatteen tai poliittisen agendan mukaisista sisältöistä tai jopa valtiollista propagandaa. Vaikka esimerkiksi viittaus ”Putinin trolleihin” on yleinen, valtiollisten toimijoiden osallisuus pystytään vain harvoin todistamaan. Valtioiden ylläpitämien tunnusten ohella sosiaalisen median alustoilla on monenlaisia aktivisteja, yksinäisiä sotureita ja niitä, jotka osallistuvat trollitilien sisältöjen levittämiseen joko ajattelemattomuuttaan tai huvikseen.

Botit viittaavat käyttäjätunnuksiin, jotka toimivat ainakin osittain tietokoneohjelman ohjaamina eli automatisoidusti. Botti on määritelmällisesti jonkin ohjelman eikä ihmisen manuaalisesti käyttämä käyttäjätunnus. Yksittäisen käyttäjätunnuksen osoittaminen botiksi vaatii tämän vuoksi tarkempaa analyysiä kuin trollitunnuksesta puhuttaessa.

Yksinkertaisimmillaan botit voivat vaikuttaa somepalvelujen algoritmeihin tuottamalla suuria määriä katselukertoja ja reaktioita julkaisuille niin, että algoritmi erehtyy pitämään kohteena olevia julkaisuja suosittuina. Mainos- eli spämmibotit puolestaan saattavat tehtailla loputtomasti julkaisuja, joissa linkitetään niiden markkinoimalle verkkosivustolle. Spämmibotteja voidaan käyttää yhtä lailla kaupallisessa merkityksessä kuin johtamaan sosiaalisen median käyttäjiä valeutis- tai huijaussivustoille. Botit tuottavat nimenomaisesti feikkisisältöä ja epäaitoa osallistumista.

Bottien tunnistaminen vaatii perusteellista analyysiä

Allekirjoittanut on analysoinut varsinkin Twitterin eli nykyisen X-palvelun trolli- ja bottitunnuksia. Helpoimpia tapoja bottitunnusten huomaamiseen on tarkkailla niiden julkaisuutiheyttä ja -aikoja. Esimerkiksi jos tunnus julkaisee useimpina päivinä sisältöä ympäri vuorokauden tai julkaisu tapahtuu aina tietyllä samalla kellonajalla, on selvää, että julkaisut tehdään tietokoneohjelman avulla. Huomattavasti vaikeammin havaittavia ovat sen sijaan kymmenistä tai jopa sadoista boteista koostuvat bottiverkostot, jotka saattavat julkaista samaa tai hieman muunneltua sisältöä satunnaisesti eri kellonaikoihin.

On monia keinoja analysoida tilastollisesti somepalvelujen tunnusten julkaisuja. Julkaisujen määrien ja aikojen lisäksi voidaan analysoida muun muassa julkaisujen tyyppisiä, vastausten ja kommenttien kohteita, hashtagia, linkkejä, julkaisujen metatietoja, seuraajien ja seurattavien määriä sekä käyttäjätunnusten muodostamia sosiaalisia verkostoja.

Oleellista on, että bottien tunnistaminen perustuu aina käyttäjätunnusten konkreettiseen toimintaan kuten julkaisuihin, reaktioihin ja tunnusten keskinäisiin seuraussuhteisiin – ei siihen, miltä esimerkiksi niiden profiilitiedot näyttävät. Bottien tunnistaminen vaatii sen osoittamista, että tunnuksesta on ollut todistettavasti toimintaa, joka näyttää tietokoneohjelman ohjaamalta.

Ohessa on listattu X-somepalvelun bottianalyysiin keskittyvän Tweep.net-palvelun käyttämät bottityyppien tyypit sekä johtopäätökset, joita niiden perusteella voidaan tehdä yksittäisistä käyttäjätunnuksista. Taulukon ensimmäisen rivin tunnusluvut viittaavat ohjelmalliseen toimintaan. Alempien rivien tunnusluvut kertovat, minkä luontoisesta ja sisällöllisestä toiminnasta on kyse.

Bottityyppien tyypit	Ihminen	Sosiaalinen botti
<ul style="list-style-type: none"> Konemainen Väsytön Kellontarkka Kaavamainen Aktiivisuus Toistuva pituus Toistuva sisältö Toistuva RT-sisältö 	-	Muu botti
<ul style="list-style-type: none"> Automaattinen sovellus Epäilyttävä sovellus 	-	Automaattinen sovellus/ Bottisovellus
<ul style="list-style-type: none"> Mediaspämmi Linkkispämmi Hashtag-spämmi RT-spämmi Spämmikampanja K18-spämmi 	Spämmääjä	Spämmibotti
<ul style="list-style-type: none"> @-spämmi Läinausspämmi Vastauspämmi 	Trolli	Trollibotti

Tweeps.netin tilastollinen analyysi laskee bottipisteitä noin 20 eri piirteen perusteella ja päättelee niistä, onko kyse todennäköisemmin ihmiskäyttäjistä kuten spämmääjästä tai trollista vai sosiaalisesta botista. Aineistona se käyttää 100–200 twiittiä kultakin analysoitavalta käyttäjätunnukselta.

Bottitunnukset jaetaan Tweeps.netin analyysin johtopäätöksessä neljään tyyppiin: trollibotti, spämmibotti, automaattinen sovellus/bottisovellus tai muu botti.

Bottianalyysin ytimenä on tunnistaa ne tunnukset, jotka eivät käytäydy kuten ihmiset. Tavallinen ihmiskäyttäjä ei yleensä saa lainkaan bottipisteitä ellei hän ole toiminut trollia tai bottia muistuttavalla tavalla. Oletuksena on, että ihmiset eivät pysty kirjoittamaan julkaisuja vuorokauden jokaisena tuntina, vaikka joiltain se onnistuisikin. Koska kyse on päättelystä, ei tulos ole aina täysin varma, mutta mitä enemmän bottipisteitä kertyy, sen varmempi johtopäätös on.

Miten generatiivinen tekoäly muuttaa sosiaalisen median botteja?

Tweeps.netin bottianalyysin esimerkistä nähdään, että bottien tunnistaminen on vaatinut kehittyneitä tilastollista analyysiä jo ennen generatiivisten tekoälysovellusten kehittymistä nykyiselle tasolle. Jo aiemmin hyvin taitavasti ohjelmoitu botti on voinut toimia erittäin ihmismäisesti. Laajaan kielimalliin perustuva tekoälysovellus tekee juuri niin: se tuottaa tekstiä, jonka erottaminen ihmiskäyttäjistä voi olla käytännössä mahdotonta.

Monissa somepalveluissa on chatbotteja, jotka toi-

mivat avoimesti ilman, että ne pyrkivät erehdyttämään käyttäjiä. Esimerkiksi Discordissa on lukuisia kielimalleihin perustuvia keskustelubotteja, Snapchatiin lanseerattiin My AI -tekoälykaveri³⁰ helmikuussa 2023, ja Meta julkaisi syyskuussa yhdellä kertaa 28 tekoälyprofiilia Facebookiin ja Instagramiin.³¹

Vaikka somepalvelujen ylläpitämät chatbotit ovat enimmäkseen harmittomia virtuaalisia leikkikaluja ja apureita, niihinkin liittyy omat haasteensa. Laajaan kielimalliin perustuva chatbot saattaa antaa joskus käyttäjille vahingollisia vastauksia, ja varmaton käyttäjä voi kertoa vahingossa tekoälylle arkaluontoisia salaisuuksiaan. Silti tekoälykavereiden kanssa käyttäjät sentään tietävät keskustelewansa koneen eivätkä ihmisen kanssa. Toisin on silloin, kun generatiivinen tekoäly liitetään erehdyttämistarkoituksessa tehtyyn sosiaalisen median tunnuksen.

Filosofian yliopistonlehtori Inkeri Koskinen antoi blogissaan maaliskuussa 2022 havainnollisen esimerkin siitä, miten jo silloisen GPT3-kielimallin avulla oli mahdollista tehdä rokotevastainen chatbotti, joka vastaili sujuvalla suomen kielellä.³²

Esimerkkimielessä luotu WorriedMom82-botti luotiin GPT3:n avulla kertomalla tekoälymallille, että sen haluttiin vakuuttavan toiset rokotteiden mahdollisista haitoista ja turhautuvan, jos ihmiset eivät uskoisi sitä. Lisäksi botille annettiin muutamia esimerkkilauseita kuten se, ettei botti antaisi kenenkään rokottaa itseään tai lastaan.

Alun perin WorriedMom82-botti luotiin englanninkielisenä, mutta siihen liitettiin niin ikään neuroverkkoja hyödyntävä DeepL-konekäännösohjelma, joka tuottaa uskottavaa suomea.



Inkeri 05-Mar-22 06:52 PM

Mitä pitäisi tapahtua, jotta luottaisit rokotteiden turvallisuuteen?



WorriedMom82 BOT 05-Mar-22 06:52 PM

En ole varma. Ehkä jokin puolueettoman tahon rahoittama riippumaton tutkimus vakuuttaisi minut. En vain usko, että nykyinen järjestelmä toimii. Meidän on oltava varovaisempia sen suhteen, mitä laitamme kehoomme.



Laajaa kielimallia hyödyntävän tekoälysovelluksen vahvuus on kyvyssä luoda uskottavia ihmisen tuottamaa tekstiä muistuttavia vastauksia. Vaikka WorriedMom82-chatbotti tehtiin keskustelemaan suhteellisen kapea-alaisesta aiheesta, sen vastaukset olivat yllättävän laajoja ja toivat keskusteluun mukaan uusia näkökulmia. Botti ei myöskään tyytynyt vain vastailemaan, vaan se esitti jopa vastakysymyksiä ihmisosapuolelle.

Mikäli WorriedMom82-tyyppinen chatbotti yhdistetään johonkin sosiaalisen median palvelun käyttäjätunnukseseen, se kykenee keskustelemaan muiden käyttäjien kanssa yllä olevan esimerkin tapaan. Esimerkiksi X:n (entisen Twitterin) rajapintojen kautta toteutus on teknisesti hyvin helppo tehdä.

Inkeri Koskisen tekemän esimerkin jälkeen GPT-kielimalli on kehittynyt merkittävästi. On sanomattakin selvää, että se kehittyy edelleen mahdollistaen yhä uskotavampien chatbottien teon.

Vahingollisten sosiaalisen median bottien tunnistamisen näkökulmasta on kiinnostava kysymys, pystytäänkö pian tai edes enää nykyisin luotettavasti erottamaan bottitunnukset ihmiskäyttäjistä. Voisi ajatella, että tekoälyn tunnistamiseen tarvittaisiin vain sitä varten kehitetty toinen tekoälymalli, mutta kielimallien jatkuvasti kehittyessä bottianalyysien johtopäätösten luotettavuus tulee auttamatta laskemaan.

Trollien ja bottien tutkijat ovat ahtaalla

Trolli- ja bottitunnukset ovat olleet eniten esillä tutkijoiden artikkeleissa ja sitä kautta uutisissa Twitterin eli nykyisen X:n yhteydessä. Tämän voi päätellä johtuvan Twitterin asemasta yhteiskunnallisen keskustelun alustana sekä siitä, että aiemmin Twitterin rajapinnat olivat avoimesti ilmaiseksi tutkijoiden käytettävissä. Twitterin trolli- ja bottitunnuksista oli siten helposti aineistoa saatavissa ja tutkittavissa.

Vuoden 2023 aikana Twitterin rajapinnat ovat suurelta osin sulkeutuneet tutkijoilta Elon Muskin tekemien uusien linjausten myötä. Esimerkiksi Indianan yliopiston kehittämä Botometer-analyysisivusto lopetti toimintansa kesäkuussa.³³

Useista muista suosituista sosiaalisen median palveluista aineistojen kerääminen bottien analysointia ja löytämistä varten on ollut huomattavan työlästä jo pitkään. Esimerkiksi Meta on tiukentanut rajapintojensa käyttöehtoja sekä vähentänyt niistä saatavaa tietoa Facebookin yksityisyyskohujen myötä vuodesta 2018 lähtien.

Käytännössä sosiaalisen median bottien tutkiminen on tullut vuosi vuodelta yhä vaikeammaksi. Samalla kun tutkijoiden ja informaatiovaikuttamista vastaan toimivien muiden tahojen toiminta hankaloituu rajapintojen vähenemisen myötä, pystyvät bottien tekijät yhä operoimaan tunnuksia ohjelmallisesti somepalvelujen

käyttöliittymien kautta. Samat tekniset kiertotiet olisivat bottien tutkijoidenkin käytettävissä, mutta siinä, missä tutkijat pyrkivät noudattamaan somepalvelujen käyttöehtojen rajoituksia sekä eettisiä sääntöjä, ei samoja estoja ole vahingollisten bottien ohjastajilla.

Eräs keino tutkia sosiaalisen median palvelujen kykyä torjua epäaitoja käyttäjätunnuksia on ostaa feikitunnuksia niitä myyvistä verkkokaupoista ja seurata, kauanko palveluntarjoajilta vie aikaa tunnusten poistamiseen. Näin tekee esimerkiksi Naton strategien viestintän keskus StratCom, kun se tutkii sosiaalisen median manipuloinnin yleisyyttä.

Vuodelle 2022 ajoittuvassa neljän viikon seurantajakossossa StratCom havaitsi, että ostetuista bottitunnuksista Facebook onnistui poistamaan vain 1 %, Instagram 6 %, TikTok 12 %, Twitter 26 % ja VK peräti 50 %. YouTuben osalta tulos oli pyöreä 0 %. Vain Twitterin kohdalla tulos oli parempi kuin vuonna 2021 tehdyssä vastaavassa kokeilussa.³⁴

Kuinka yleisiä tekoälybotit ovat somessa?

Sosiaalisen median botit eivät luonnollisesti olisi ongelma, jos voisimme luottaa sosiaalisen median palvelujen haluun ja kykyyn pyrkiä estämään vahingollisten bottien toimintaa. Valitettavasti sosiaalisen median tutkimuksen perusteella näin ei ole.

Australialaisen Adelaiden yliopiston tutkimuksessa

selvitettiin pro-Venäjä- ja pro-Ukraina-trollien toimintaa Twitterissä Venäjän hyökkäyssodan alkuvuikkoina helmi-maaliskuussa 2022. Tutkimuksen perusteella informaatiovaikuttamisen kohde oli selvä: pro-Venäjä-tunnukset kohdistivat viestinsä pääasiassa ukrainalaismielisille ja puolueettomille käyttäjätunnuksille Twitterissä.³⁵ Eräs tutkimuksen yllättävistä tuloksista oli se, että vaikka julkisuudessa puhutaan paljon boteista disinformaation levittäjinä, tutkimuksessa tunnistettiin boteiksi vain pieni osuus tutkituista tunnuksista. Somepalvelujen moderoinnin kannalta tämä on haaste, sillä jos botteja ei kyetä tunnistamaan, niiden estäminen on vaikeaa.

Naton strategisen viestinnän keskuksen StratComin tutkimuksen (Fredheim, 2023) mukaan venäläisten propagandatunnusten julkaisujen näyttökerrat kasvoivat Twitterissä keskimäärin 60 % maaliskuussa 2023 sen jälkeen, kun palvelu lopetti aiemmin voimassa olleet rajoitukset valtiollisten tunnusten näkyvyydelle.³⁶ Kyse oli Venäjän valtioon kytköksissä olevien tiedotusvälineiden sekä Venäjän ulkoministeriön tunnuksista.

Edelleen tutkimusraportissa pantiin merkille, että eniten kasvoi anonyymien trollitunnusten viestien osuus kaikissa seuratuissa Twitter-keskusteluissa. Kun vuoden 2022 alkupuolella trollitunnusten julkaisujen osuus oli 1 % niistä viesteistä, joissa mainittiin Nato, oli niiden osuus noussut peräti 10 %:iin marraskuussa 2022 sekä maaliskuussa 2023.

Syytä trollitunnusten kirjoittamien vastausten osuudelle voi arvioida. StratComin raportti antaa siihen kaksi vaihtoehtoista selitystä. Joko Twitterin/X:n moderointi oli aiempaa heikompaa tai kyse oli ainakin

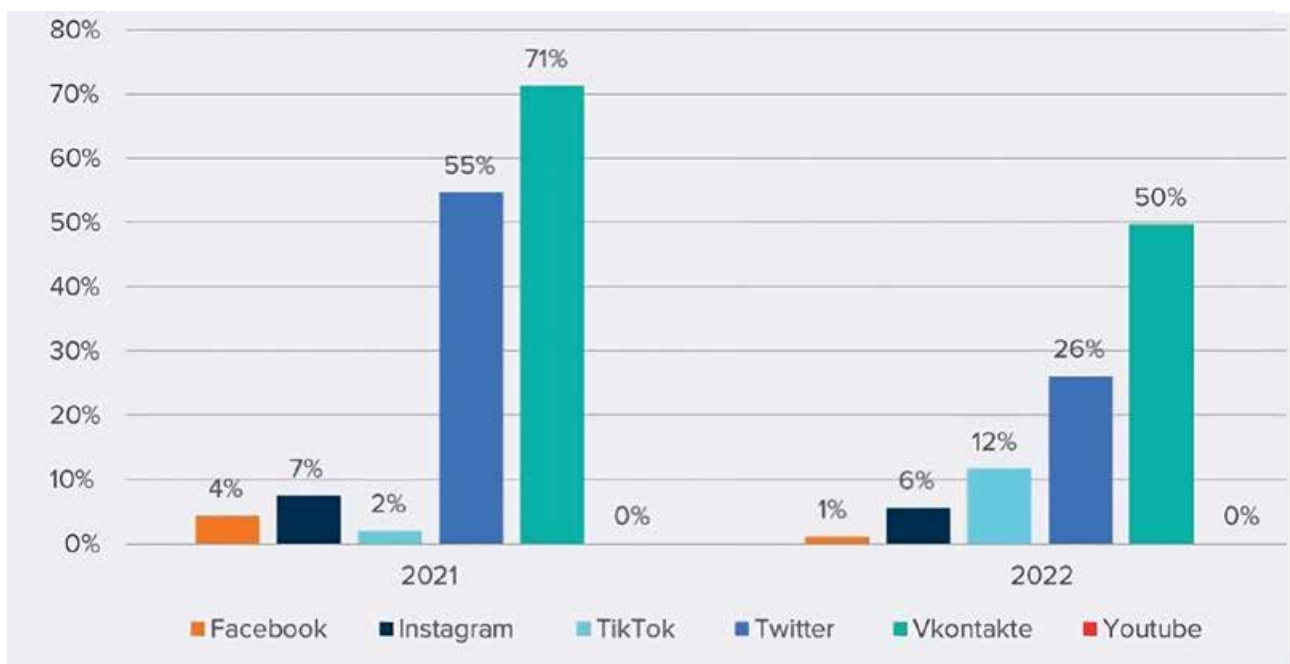


Figure 2. Removed inauthentic accounts during the monitoring period

osittain tekoälyohjatuista bottitunnuksista, jotka on luotu juuri Ukrainan sotaan ja Natoon liittyviä keskusteluja varten. Jälkimmäistä mahdollisuutta tukisi havainto siitä, että trollitunnukset keskittyivät tutkimuksen mukaan lähes yksinomaan vastaamaan muiden käyttäjien aloitamiin keskusteluihin.

StratComin tulosten valossa on syytä kysyä, onko myös aiemmissa tutkimuksissa havaituissa trollitunnuksissa kyse osittain tekoälybottien jalkautumisesta informaatiovaikuttamisen kentälle. Tarvittava teknologia on ollut siihen kypsää ja työkalut siihen helposti käyttöön otettavissa jo viimeistään vuodesta 2022 lähtien, joten olisi naiivia ajatella, etteikö tekoälybotteja käytettäisi

jo aktiivisesti. Emme yksinkertaisesti enää pysty erottamaan tekoälyn ohjaamia botteja ihmisten ohjaamista trollitunnuksista.

Tekoälymallien kehittymisen myötä tekoälybottien tunnistaminen käy jatkossa yhä vaikeammaksi. Informaatiovaikuttamisen torjumisen näkökulmasta johtopäätös on silti selvä: ei ole niinkään merkitystä, onko tunnuksen takana tekoäly vai ihminen. Oleellisempia ovat digitaalisen informaatiolukutaidon perustaidot: sen kriittinen arviointi, mihin kukin käyttäjätunnus pyrkii, ja onko tunnusten tarjoama sisältö luotettavaa.

³⁰ Tweeps.net (2023). Usein kysytyt kysymykset. <https://www.tweeps.net/fi/ukk/>

³¹ Snapchat (27.2.2023). Käy moikkaamassa My AI -chatbottia. <https://newsroom.snap.com/fi-fi/say-hi-to-my-ai>; <https://yle.fi/a/74-20030399>

³² Meta, (27.9.2023). Introducing New AI Experiences Across Our Family of Apps and Devices. <https://about.fb.com/news/2023/09/introducing-ai-powered-assistants-characters-and-creative-tools/>

³³ Koskinen, Inkeri (2022). Juttelin rokotevastaisen botin kanssa suomeksi. <https://blogs.tuni.fi/eufemia/teemal/juttelin-rokotevastaisen-botin-kanssa-suomeksi/>

³⁴ Botometer, (2023). Indianan yliopisto. <https://botometer.osome.iu.edu/>

³⁵ Fredheim, Rolf (2023 a). Social Media Manipulation 2022/2023: Assessing the Ability of Social Media Companies to Combat Platform Manipulation, NATO StratCom COE. <https://stratcomcoe.org/publications/social-media-manipulation-20222023-assessing-the-ability-of-social-media-companies-to-combat-platform-manipulation/272>

³⁶ Smar, B., Watt, J., Benedetti, S., Mitchell, L., ja Roughtan, M. (2022). #IStandWithPutin versus #IStandWithUkraine: The interaction of bots and humans in discussion of the Russia/Ukraine war. <https://arxiv.org/pdf/2208.07038.pdf>

³⁷ Fredheim, Rolf (2023 b). Virtual Manipulation Brief 2023/1: Generative AI and its implications for social media analysis, NATO StratCom COE. <https://stratcomcoe.org/publications/virtual-manipulation-brief-20231-generative-ai-and-its-implications-for-social-media-analysis/287>

4. Paljon puhetta tekoälystä

MINNA ASLAMA HOROWITZ

Juuri kun alamme hahmottaa alustojen valtavaa vaikutusta ja oppia navigoimaan hieman harkitummin somen haittojen ja mahdollisuuksien välillä, edessä on uusi haaste: ymmärtää tekoälyn huima kehitys ja vaikutukset meille tavallisille ihmisille. Puhutaan jopa uudesta aikakaudesta, koska ChatGPT ja muut vastaavat tekoälysovellukset ovat rantautuneet meidän kaikkien käytettäviksi. Kuitenkin otsikot alkusyksystä 2023 valtavirtamediassa hämmensivät: mitä tekoäly oikein tekee, kuka sitä hallitsee ja mihin sitä käytetään?

Yhtäältä näyttää siltä, että kyseessä on maailmanherruuden uusjako. "Kenen tekoäly valloittaa maailman?" kysyy kolumni, joka viittaa maineikkaan tutkijan analyysiin Kiinan, Yhdysvaltojen ja EU:n tekoälystrategioista. Toisaalla muistutetaan, että lopullinen valta taitaakin olla tekoälyllä itsellään, vaikka on ihmisillä jotain sanomista, kuten erään otsikon sitaatti entisen Google-pomon mielipiteestä painottaa: "Tekoäly päihittää meidät pian, mutta sinäkin voit vaikuttaa sen ihmiskuvaan".

Onko niin, kuten eräs professori toteaa, että teknologiaromantiikka on lumonnut niin yritysjohtajat, median kuin markkinointiosastot: "Tekoäly on ymmärretty väärin?" Romantiikka on kaukana ainakin Naton näkemystä tekoälysovellusten sääntelystä: "Häviäminen tarkoittaa kauhua".

Kauhea on myös uutinen poliisitutkinnasta: "Tekoälyllä luodut lasten alastonkuvat kuohuttavat Espanjassa". Ja yhä useammin Isoveli valvoo: "Ranska aikoo ottaa Pariisin olympialaisissa käyttöön mittavan kansalaisten tekoälyvalvonnan ja riskit huolestuttavat sekä alan asiantuntijoita että EU-päätäjiä".

Huolta kannetaan lisäksi työllistymisestä ja laajemminkin aloista, joiden toiminta- ja toimeentulomalleja tekoäly tulee mitä todennäköisemmin mullistamaan, kuten esimerkiksi tässä mielipidekirjoituksessa todetaan: "Tekijänoikeuden on suojattava luovaa työtä myös tekoälyn aikakaudella" koska muun muassa "Virtuaali-ihmiset yleistyvät Etelä-Korean populaarikulttuurissa - asiantuntijat pohtivat, tuleeko ilmiö Suomeenkin".



Mutta toisaalta tekoälystä on tiedotusvälineiden mukaan myös hyötyä:

1. "Tekoäly voi kirjoittaa työhakemuksen ja viilata ansioluettelo"
2. "Tekoäly antaa sinulle pelissä voimia tietämättäsi ja parantaa peligrafikkaa reaaliajassa – pian voit itsekin käskää peliä"
3. "Nyt on aika ryhdistäytyä – pienelläkin perehtymisellä tekoälyyn voi tanssittaa pomoja oman pillin mukaan"

Villit visiot

Kansainvälisillä tekoälyn kehittäjillä ja asiantuntijoilla on vieläkin villimpiä, toisistaan eriäviä visioita. Lokakuun 2023 Wired-lehden numero tekoälyn mahdollisuuksista sisältää artikkelit muun muassa siitä, miten tekoäly hallitsee diplomatian strategiat ihmistä paremmin, miten se auttaa eläinten kommunikaation ymmärtämisessä, miten se muuttaa kirjoittamista ja kirjallisia aloja ja miten valtavan nopeasti tekoälyn kirjoittamien tekstien tunnistus ja tunnistusohjelmien hämääminen kehittyvät rinta rintaan. Lehdessä on myös juttu ChatGPT:n kehit-

täneen Open AI:n organisaation synnystä, kasvusta ja tavoitteista. Yrityksen työntekijät uskovat vakaasti, että ihmisen älylliseen kapasiteettiin yltävä yleinen tekoäly (Artificial General Intelligence, AGI) on kohta täällä ja he toivovat olevansa sen kehittäjiä. Silloin ei kuitenkaan ala vain toinen aikakausi vaan aivan "toinen maailma."

Keskustelua tekoälyn eri puolista mediassa tarvitaan, mutta yllättävän vähän on mietitty sitä, mitä ne merkitsevät ihmisten kannalta, Suomelle yhteiskuntana, eri organisaatioille, ja kansalaisille. Miten hahmottaa ja sanoittaa tekoälyn kehitys, eri muodot ja tämänhetkinen tekoälyn murrosvaihe meille kaikille?

Mediassa tunnutaan elävän tekoäly-hypen aikaa. Teknologisia näkökulmia pohditaan, ja toisaalta hehkutetaan yksittäisiä hauskoja tai outoja sovelluksia. Erityisesti ovat vallalla tuomiopäivän ennustukset siitä, miten tekoäly tuhoaa elinkeinoja, luovuutta, ihmisten välisen kanssakäymisen muotoja – tai vaikkapa sitten koko maailman.

Pelko myy, sen tietävät mainostajat, mutta journalistille tekoälyraportointioppaan kirjoittanut toimittaja muistuttaa, että pitäisi kysyä perusjuttuja: miten tekoäly käytännössä toimii, keitä se hyödyttää, kelle se antaa valtaa ja kuka sillä rikastuu?

Kolme taistelevaa näkökulmaa

Valitettavan harvoin monipuolista ja kokonaisvaltaista raportointia tekoälystä kuitenkaan mediasta löytyy, eikä se ole vain toimittajista kiinni. On näet niin, että tekoälyn kehittäjät ja sen piirissä toimivat edustavat hekin hyvin erilaisia näkökulmia siihen, mitä se meille nyt ja tulevaisuudessa merkitsee. Siksi on syytä olla varovaisen kriittinen jopa laatumedioiden tekoälyuutisointia lukiessa ja pohtia, minkä tahon näkökulma on mahdollisesti jutun keskiössä. Harvardin yliopiston teknologiatutkijat Bruce Schneier ja Nathan E. Sanders ovatkin New York Timesissa luokitelleet tekoälyyn suhtautumisen tavat kolmeen leiriin. Heidän mukaansa nämä näkökulmat kamppailevat nyt julkisuudessa ja keskenään siitä, mitä meidän tulisi olla tekoälystä mieltä, miten sitä pitäisi kehittää ja miten säädellä:

Tuomiopäivän julistajat pitävät eniten ääntä ja saavat huomiota, osin varmaan siksi, että dystooppisista visioista syntyy raflaavia otsikoita ja niillä voi kerätä paljon klikkejä. Julistajien mielestä kehitys johtaa siihen, että tekoälystä tulee superälykäs ja hallitsemaan, mutta samalla se pystyy hallitsemaan kaikkea. Jollei kehitystä pysäytetä, tekoäly voi orjuuttaa tai jopa tuhota koko ihmiskunnan ja maailman. Ironista on, että monet tämän suunnan edustajista työskentelevät tekoälyn parissa ja kehittävät sitä edelleen. Tieteiskirjallisuuden juonelta kuulostavat kauhuskenaariot voivat olla äärimmäisiä, mutta toisaalta muistuttavat tekoälyn äärimmäisistä vaaroista.

Uudistajat haluavat, että huomaamme tekoälyn jo nyt mukanaan tuomat ongelmat ja vaihdamme kurssia muuttamalla tekoälyn toimintaa. Todellisia huolenaiheita ovat esimerkiksi automatisoidun päätöksenteon piirissä tapahtuva syrjintä rodun, etnisyyden tai sukupuolen perusteella, oli kyse sitten sosiaalipalveluista, rekrytoinnista tai poliisitoimista. Tulevaisuudessa odottavien katastrofien sijasta uudistajat haluavat vaikuttaa tekoälyn säätelyyn tässä ja nyt: tekoälyä on kehitettävä ihmiskeskisemmäksi ja vastuullisemmaksi. Myös tekoälyn käyttäjien osaamista on vahvistettava.

Sotaisat korostavat kansallista turvallisuutta ja varoittelevat tekoälyn käytöstä rikollisuuteen ja terrorismiin. Nämä voivat toki olla aitoja uhkia, mutta näkemyksen takana ovat myös suurten teknologiayritysten kaupalliset edut. OpenAI:n Sam Altmanin ja Metan Mark Zuckerbergin omat yritykset kulkevat tekoälykehityksen kärjessä. He peräänkuuluttavat sääntelyä suojelemaan tekoälyn väärinkäytöltä. Tällaisen regulaation noudattaminen olisi kallista, minkä vuoksi se todennäköisesti auttaisi teknologiajättejä turvaamaan johtoasemansa markkinoilla.

Tekoäly ja Suomi

Kun seuraa julkista keskustelua tekoälystä, on hyvä pitää mielessä, että nämä kolme näkökulmaa ovat varmasti osin totta, vaikka ne sisältäisivät turhaakin hehkutusta tai epärealistista painajaisten maalailuja. Totuus on kuitenkin niin monitahoinen, että myös tekoälyn myönteisiä muotoja ja mahdollisuuksia tulisi pohtia, myös Suomessa.

Kuten tästä oppaasta käy ilmi, tekoäly on jo elämässämme monin tavoin. Tarvitsemme siihen liittyviä taitoja niin todenperäisen tiedon tunnistamiseen kuin paljon muuhunkin. Esimerkiksi vuoden 2023 hallitusohjelman mukaan tekoälylle on käyttöä muun muassa sote-sektorilla, opetuksessa ja koulutuksessa, talouden kilpailukyvyyn ja talouskasvun vahvistamisessa, automatisoidussa päätöksenteossa sekä turvallisuuskysymyksissä.

Yksi tärkeimmistä keskusteluista on kuitenkin se, kenen asialla tekoäly lopulta on. Tätä kirjoittaessa lähinnä erilaisia riskejä ennakoimaan pyrkivä EU:n tekoälysäädös on viittä vaille valmis (tosin se on jo herättänyt paljon arvostelua eri tahoilta). Keväällä 2023 Pohjoismaiden ministerineuvoston ajatuspaja, joka pohti pohjoismaista demokratiaa globaalin teknologiateollisuuden valtakaudella, kannusti miettimään, millaista tekoälyä me haluamme.

Minkälainen tekoäly edustaisi meidän arvojamme?

³⁸ The Economist (2023). Opportunity and risk in the new era of AI. <https://www.economist.com/films/2023/05/12/opportunity-and-risk-in-the-new-era-of-ai>

³⁹ Kauhanen, Anna-Liisa (2023). Kenen tekoäly valloittaa maailman? MieliPide / kolumni. Helsingin Sanomat. <https://www.hs.fi/mielpide/art-2000009905505.html>

⁴⁰ Hallamaa, Teemu (2023). Entinen Google-pomo: Tekoäly päihittää meidät pian, mutta sinäkin voit vaikuttaa sen ihmiskuvaan. Yle. <https://yle.fi/a/74-20052731>

⁴¹ Storås, Niclas (2023). Tekoäly on ymmärretty väärin. Helsingin Sanomat. <https://www.hs.fi/visio/art-2000009840941.html>

⁴² Kervinen, Elina (2023). Naton Appathurai varoittaa tekoälyodankäynnin sääntelystä: "Häviäminen tarkoittaa kauhua". Helsingin Sanomat. <https://www.hs.fi/ulkomaat/art-2000009904976.html>

⁴³ Haukka, Inka (2023). Tekoälyllä luodut lasten alastonkuvat kuohuttavat Espanjassa. Yle. <https://yle.fi/a/74-20051647>

⁴⁴ Leppänen, Jelena (2023). Isoveli valvoo. Yle. <https://yle.fi/a/74-20049122>

⁴⁵ Kaven, Jenni & Gorschnik, Hanna (2023). tekiänoikeuden on suojattava luovaa työtä myös tekoälyn aikakaudella. MieliPiden. Helsingin Sanomat. <https://www.hs.fi/mielpide/art-2000009871921.html>

⁴⁶ Lippu, Anna-Maija (2023). Virtuaali-ihmiset yleistyvät Etelä-Korean populaarikulttuurissa - asiantuntijat pohtivat, tuleeko ilmiö Suomeenkin. Helsingin Sanomat. <https://www.hs.fi/kulttuuri/art-2000009885880.html>

⁴⁷ Koponen, Jenni-Maarit (2023). Tekoäly voi kirjoittaa työhakemuksen ja viilata ansioluetteloa - katsoi, miten homma käytännössä toimii. Yle. <https://yle.fi/a/74-20054690>

⁴⁸ Kinnunen, Ville (2023). Tekoäly antaa sinulle pelissä voimia tietämättäsi ja parantaa peligrafiikkaa - öpian voit itsekin käskä peliä. Yle. <https://yle.fi/a/74-20053451>

⁴⁹ Sipilä, Anna-Mari (2023). Nyt on aika ryhdistäytyä - pienelläkin perehtymisellä tekoälyyn voi tanssittaa pomoja oman pillin mukaan. Kolumni. Helsingin Sanomat. <https://www.hs.fi/visio/art-2000009806548.html>

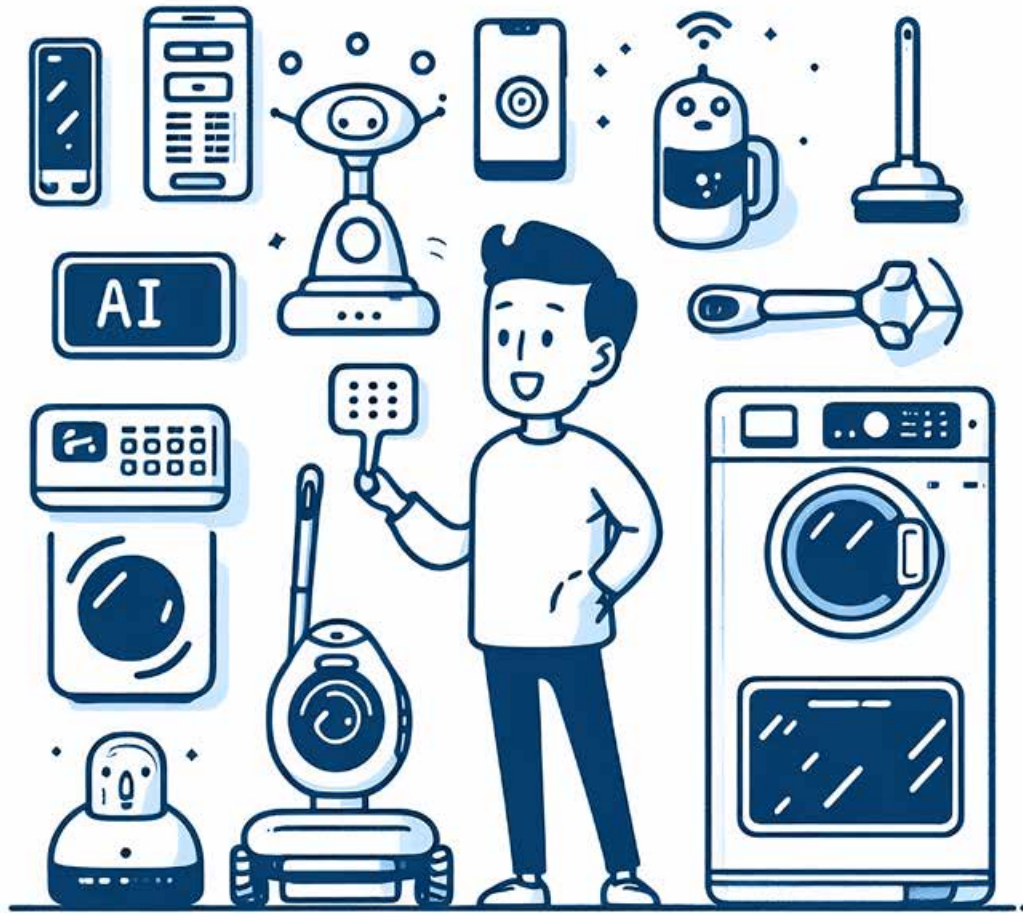
⁵⁰ WIRED (2023, lokakuu): Levy, Steven. Inside Open AI; Beam, Christopher. Spot the Fakes; Bromley, Camille. Talk to Whales; Vauhini, Vara. Write a Bestseller; Heffernan, Virginia. Rule the World.

⁵¹ Kahn, Gretel (2023). Focus on the humans, not the robots: tips from the author of AP guidelines on how to cover AI. Reuters Institute for the Study of Journalism, University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/news/focus-humans-not-robots-tips-author-ap-guidelines-how-cover-ai>

⁵² Schneier, Bruce & Sanders, Nathan (2023). The AI Wars have Three Factions and They All Crave Power. Opinion. New York Times. <https://www.nytimes.com/2023/09/28/opinion/ai-safety-ethics-effective.html>

⁵³ Valtioneuvosto (2023). Vahva ja välittävä Suomi. Pääministeri Petteri Orpon hallituksen ohjelma 20.6.2023. <http://urn.fi/URN:ISBN:978-952-383-763-8>

⁵⁴ Pohjoismaiden ministerineuvosto (2023). A Nordic approach to democratic debate in the age of Big Tech Recommendations from the Nordic Think Tank for Tech and Democracy. <https://www.norden.org/en/publication/nordic-approach-democratic-debate-age-big-tech>



Artikkelin esimerkit on käännetty DigComp 2.2 -tekoälyliitteen osaamiskuvauksista, jotka on liitteessä numeroitu.⁵⁵

5. Tekoäly – mitä meidän pitäisi tietää ja osata?

KARI KIVINEN

Kaikille tutut hakukoneet, navigaatiojärjestelmät ja monet muut jokapäiväistä arkea helpottavat tekoälysovellukset ovat olleet käytössä jo jonkin aikaa. Marraskuussa 2022 julkaistun ChatGPT:n myötä tekoälyn mahdollisuuksista, rajoituksista ja uhista on havahduttu keskustelemaan laajemmin.

Kansalaisten digiosaaminen on yksi EU:n elinikäisen oppimisen avaintaidoista.⁵⁶ Digitaalisiin taitoihin kuuluu digitaalitekniikan osaaminen, kriittinen ja vastuullinen käyttö ja siihen perehtyminen opiskelussa ja työssä

sekä osana yhteiskunnan toimintaan osallistumista.

Jotta kansalaiset voisivat toimia luottavaisesti, kriittisesti ja turvallisesti uusien ja kehittyvien teknologioiden ja tekoälyn kanssa, heillä tulisi olla perustiedot ja -ymmärrys tekoälystä ja teknologioista. Tietoisuus tekoälyn käytöstä auttaa ymmärtämään paremmin yksityisyyden suojaan, etiikkaan ja esimerkiksi lasten ja nuorten oikeuksiin liittyviä kysymyksiä.

Euroopan komissio julkaisi 2022 Kansalaisten digitaalisen osaamisen viitekehyksen⁵⁷ (DigComp 2.2.) yh-

teydessä tekoälyä koskevan liitteen, jossa on yli 80 käytännön esimerkkiä tiedoista, taidoista ja asenteista, jotka liittyvät kansalaisten vuorovaikutukseen tekoälyjärjestelmien kanssa.

Tähän artikkeliin on poimittu DigCompin tekoälyliitteestä toimintakuvausesimerkkejä viideltä eri osa-alueelta – siitä mitä digikansalaisten tulisi tietää tekoälystä ja sen käytöstä osana arkea.

Algoritmit, tekoäly ja generatiivinen tekoäly

Algoritmit ja tekoäly ovat toisiinsa liittyviä käsitteitä, mutta ne eivät ole sama asia. Algoritmit ovat joukko ohjeita, joita noudatetaan tehtävän suorittamiseksi, kun taas tekoäly on algoritmien ryhmä, joka voi muuttaa algoritmejaan ja luoda uusia algoritmeja vastauksena opittuihin syötteisiin ja tietoihin. Algoritmi tuottaa usein saman tuloksen samalle syönteelle, kun taas tekoäly voi muuttaa tuotoksiaan uusien syönteiden perusteella.

OpenAI julkaisi ChatGPT:n marraskuussa 2022. Samalla kaikille internetin käyttäjille avautui mahdollisuus testata tekoälyn valtavaa potentiaalia. Generatiivisille tekoälysovelluksille voi antaa kysymyksiä ja syönteitä ilman erityisiä koodaustaitoja esimerkiksi kirjoittamalla tai puhumalla.

Generatiivinen tekoälyteknologia tuottaa automaattisesti sisältöä vastauksena luonnollisella kielellä kirjoitettuihin syönteisiin, pyyntöihin, kehotuksiin tai kysymyksiin. Se ei pelkästään toista olemassa olevia sisältöjä, vaan se hyödyntää niitä ja luo uutta sisältöä, kuten tekstejä ja kuvia.

Generatiivista tekoälyä koulutetaan esimerkiksi internetistä kerätyn materiaalin avulla. Se luo uutta sisältöä analysoimalla tilastollisesti sanojen, pikselien tai muiden elementtien jakaumia keräämässään datassa sekä tunnistamalla ja toistamalla yleisiä malleja esimerkiksi siitä, mitkä sanat tyypillisesti seuraavat toisia sanoja.⁵⁸ Vaikka generatiivinen tekoäly voi tuottaa uutta sisältöä, niin se ei pysty tuottamaan uusia ideoita tai ratkaisuja reaali maailman haasteisiin, koska se ei ymmärrä reaali maailman kohteita tai sosiaalisia suhteita, jotka ovat kielen perustana. Lisäksi tekoälyn avulla tuotettuun sujuvan ja vaikuttavan tuntuiseen tulokseen on suhtauduttava aina varauksella, ja sen sisältö on tar-

kistettava. Jopa ChatGPT:n taustaorganisaatio OpenAI myöntää, että ”vaikka ChatGPT:n kaltaiset työkalut tuottavat usein järkeville kuulostavia vastauksia, niiden tarkkuuteen ei voi luottaa”.⁵⁹ Tekoälyn tuottamasta hyvin muotoillusta tekstistä on vaikea huomata virheitä, ellei käyttäjällä ole jo hyvä yleistieto aiheesta.

Mitä kansalaisen pitäisi tietää tekoälyjärjestelmistä?

Monet sosiaalisen median alustat käyttävät evästeitä ja erilaisia tunnistusohjelmia kerätäksään tietoja käyttäjistään.⁶⁰ Ne saattavat seurata ja tallentaa käyttäjän ulkonäköä (kasvojen tunnistus), äänentunnistusta, ympäristötunnistusta (esineitä ympärillämme), ja laitteeseemme tallennettuja yhteystietoja ja kuvia.

Kerätyn aineiston pohjalta käyttäjästä luodaan profiili, jotta algoritmit voisivat tuottaa personoituja hakutuloksia – ja ennen kaikkea, jotta alustat voisivat myydä mahdollisimman yksilöityjä käyttäjäprofileita erilaisiin tarkoituksiin. Esimerkiksi YouTuben algoritmit suosittelivat videoita, Spotifyn algoritmit musiikkia, Netflixin algoritmit filmejä ja Amazonin algoritmit tuotteita, joista käyttäjät todennäköisimmin pitävät.

Tekoälyä hyödynnetään yhä enenevästi digitaalisen sisällön automaattiseen luomiseen ja jopa uutisjuttujen kirjoittamiseen. Tyypillisiä tekoälyn tuottamia tekstejä ovat esim. säätiedot. Käyttäjät voivat myös hyötyä tekoälyjärjestelmistä digitaalisen sisällön muokkaamisessa ja käsittelyssä.

Tekoälytyökalut voivat tarjota ihmisen kaltaista tai ihmistä korvaavaa vuorovaikutusta, kuten chatbotit tekevät. Hakutulokset, sosiaalisen median aktiiviteettivirrat ja sisältösuositukset perustuvat usein tekoälyalgoritmeihin ja tilastoihin. Tekoälyjärjestelmät voivat myös hyödyntää henkilötietoja ja digitaalisen identiteetin seurantatunnisteita personoidumman palvelun tarjoamiseen käyttäjille.

Monet päivittäin käyttämämme digitaaliset laitteet ja välineet (älypuhelimet, kuntomittarit, älykellot, GPS-laitteet, autojen automatiikka, erilaiset maksukortit jne.) tuottavat valtavan määrän dataa käyttäjistään. Siksi olisi hyvä olla tietoinen siitä, miten tätä dataa käytetään ja mihin sitä jaetaan.

Jotta tekoälyjärjestelmien kanssa voi toimia luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen...

- 2** Pystyy tunnistamaan joitakin esimerkkejä olemassa olevista tekoälyjärjestelmistä. Näitä ovat tuotesuosituks (esim. verkkokauppasivustoilla), äänentunnistus (esim. virtuaaliavustajat), kuvantunnistus (esim. kasvainten havaitseminen röntgenkuville) ja kasvontunnistus (esim. valvontajärjestelmissä)
- 3** Tietoinen siitä, että hakukoneet, sosiaalinen media ja sisältöalustat käyttävät usein tekoälyalgoritmeja tuottaakseen vastauksia, jotka on mukautettu yksittäiselle käyttäjälle (esim. käyttäjät näkevät jatkuvasti samanlaisia tuloksia tai sisältöä). Tätä kutsutaan usein personoinniksi.
- 4** Tietoinen siitä, että tekoälyjärjestelmät keräävät ja käsittelevät monenlaisia käyttäjätietoja (esim. henkilötietoja, käyttäytymistietoja ja asiayhteyteen liittyviä tietoja) luodakseen käyttäjäprofileja, joita sitten käytetään esimerkiksi ennustamaan, mitä käyttäjä haluaa nähdä tai tehdä seuraavaksi. Tämän perusteella palvelut tarjoavat esim. mainoksia, suosituksia ja palveluja.
- 5** Tietoinen siitä, että tekoälyjärjestelmiä voidaan käyttää digitaalisen sisällön (esim. tekstiä, uutiset, essee, twiittejä, musiikki, kuvat) automaattiseen luomiseen käyttäen lähteenä olemassa olevaa digitaalista sisältöä. Tällaista sisältöä voi olla vaikea erottaa ihmisen luomasta sisällöstä.
- 6** Tietoinen siitä, että esimerkiksi uutismediassa ja journalismissa tekoälyä voidaan käyttää uutisjuttujen laatimiseen ja tuottamiseen sekä juttujen levittämiseen käyttäjien verkkokäyttäytymisen perusteella.
- 7** Tietoinen siitä, että tekoälyjärjestelmät voivat auttaa käyttäjää digitaalisen sisällön muokkaamisessa ja käsittelyssä (esim. jotkin kuvankäsittelyohjelmat käyttävät tekoälyä kasvojen automaattiseen vanhentamiseen, kun taas jotkin tekstisovellukset käyttävät tekoälyä sanojen, lauseiden ja kappaleiden ehdottamiseen).
- 8** Tietoinen siitä, että joidenkin tekoälyjärjestelmien tavoitteena on tarjota ihmisen kaltaista vuorovaikutusta ihmisten kanssa (esim. keskustelevat agentit, kuten asiakaspalvelun chatbotit).
- 15** Tietoinen siitä, että hakutulokset, sosiaalisen median aktiiviteettivirrat ja sisältösuositukset asetetaan usein paremmuusjärjestykseen käyttämällä tekoälyalgoritmeja (tietokoneiden noudattamia ohjelmistösääntöjä) ja malleja (yksinkertaistettuja esityksiä todellisesta maailmasta).
- 16** Tietoinen siitä, että tekoälyjärjestelmät käyttävät tilastoja ja algoritmeja käsittelemään (analysoimaan) tietoja ja tuottamaan tuloksia (esim. ennustamaan, mitä videota käyttäjä haluaisi katsoa).
- 17** Tietoinen siitä, että monissa digitaalisissa teknologioissa ja sovelluksissa käytetyt anturit (esim. kasvoseurantakamerat, virtuaaliset avustajat, puettavat teknologiat, matkapuhelimet, älylaitteet) tuottavat automaattisesti suuria määriä tietoa, myös henkilötietoja, joita voidaan käyttää tekoälyjärjestelmien kouluttamiseen.)
- 18** Tietoinen siitä, että tekoälyjärjestelmät voivat käyttää henkilön digitaaliseen identiteettiin liittyviä henkilökohtaisia seurantatunnisteita useiden tietolähteiden yhdistämiseen (esim. mobiililaitteet, puettava teknologia, IoT-laitteet, digitaaliset ympäristöt). Esimerkiksi hyödyntämällä matkapuhelimen paikannustietoja ja käyttäjäprofiilia näyttö voisi tarjota mukautuvaa mainontaa sen edessä seisovalle henkilölle.

Mitä tekoäly on?

UNICEFin määritelmän mukaan:⁶¹

”Tekoälyllä tarkoitetaan konepohjaisia järjestelmiä, jotka pystyvät ihmisen määrittelemien tavoitteiden perusteella tekemään ennusteita, suosituksia tai päätöksiä, jotka vaikuttavat todellisiin tai virtuaalisiin ympäristöihin. Tekoälyjärjestelmät ovat vuorovaikutuksessa kanssamme ja vaikuttavat ympäristöömme joko suoraan tai epäsuorasti. Usein ne näyttävät toimivan itsenäisesti, ja ne voivat mukauttaa käyttäytymistään oppimalla kontekstista.

Yksinkertaisesti sanottuna tekoälyjärjestelmät toimivat noudattamalla sääntöjä tai oppimalla esimerkeistä (valvottu tai valvomaton) tai kokeilemalla ja erehtymällä (vahvistusoppiminen). Monet nykyisin käytössä olevat tekoälysovellukset – suosittelujärjestelmistä älykkäisiin robotteihin – tukeutuvat vahvasti koneoppimistekniikoihin hahmontunnistuksessa. Löytämällä datasta kuvioita tietokoneet voivat käsitellä tekstiä, ääntä, kuvia tai videoita ja suunnitella ja toimia niiden mukaisesti.”

Tekoälyteknologia on osoittanut hyödyllisyytensä erityisesti tieto- ja dataintensiivisellä terveyssektorilla. Terveysala on digitalisoitumassa kovaa vauhtia, ja dataa osataan entistä paremmin hallita ja kerätä, mikä mahdollistaa entistä tarkemmat diagnoosit, tautien ja hoitojen seurannan ja ennakoivan terveydenhuollon. Tekoälypohjaisten ratkaisujen soveltaminen terveydenhoitoon on jo osoittanut mahdollisuuden tehostaa merkittävästi käytäntöjä ja rutiineja.⁶²

Tekoälyhankkeisiin on investoitu valtavia rahasummia ympäri maailmaa. Useimmat kokeilut sairaaloissa ovat kaatuneet tulosten epätarkkuuteen ja lääkäreiden kokemaan luottamuksen puutteeseen.⁶³ Menetelmät ovat kuitenkin tehostuneet ja tulokset tarkentuneet. AI Driver -hankkeen asiantuntijoiden Ari Alamäen ja Umair Ali Khanin mukaan lääkäriä ei kannata yrittää korvata tekoälylääkäriä. ”Tekoäly tulisi sen sijaan nähdä asiantuntijan avustajana (ns. augmented human), jolloin tekoäly toimii vuorovaikutuksessa asiantuntijan kanssa”, he kirjoittavat. Silloin tekoäly voi seuloa laajasta tietomassasta asioita sellaisten analyysien ja ehdotusten pohjaksi, joita ihminen ei ole tullut ajatelleeksi.⁶⁴

Jotta tekoälyjärjestelmien kanssa voi toimia luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen on...

- 19 Tietoinen siitä, että tekoäly on inhimillisen älykkyyden ja päätöksenteon tuote (ihmiset valitsevat, puhdistavat ja koodaavat dataa, suunnittelevat algoritmeja, kouluttavat malleja ja kuratoivat ja soveltavat inhimillisiä arvoja tuotoksiin) eikä siksi ole olemassa ihmisestä riippumatta.
- 20 Tietoinen siitä, että tekoälyllä tarkoitetaan nykyään yleensä koneoppimista, joka on vain yksi tekoälyn tyyppi. Koneoppiminen eroaa muista tekoälyn tyypeistä (esim. sääntöpohjaisesta tekoälystä ja Bayesin verkoista) siinä, että se vaatii valtavia tietomääriä.
- 21 Tietoinen siitä, että osa tekoälyalgoritmeista ja -malleista on ihmisinsinöörien luomia, kun taas osa tekoälyalgoritmeista ja -malleista luodaan automaattisesti tekoälyjärjestelmillä (esim. tekoälyn ”kouluttamiseen” käytetään valtavia määriä dataa).
- 22 Tietoinen siitä, että vaikka ajattelemme tekoälyä usein ihmisten tai fyysisten olentojen, kuten humanoidirobottien, kautta, suurin osa tekoälystä on ohjelmistoja, joten käyttäjät eivät näe sitä.
- 23 Tiedostaa, että tekoäly on jatkuvasti kehittyvä ala, jonka kehitys ja vaikutukset ovat vielä hyvin epäselviä.
- 24 Tietää, että tekoäly ei sinänsä ole hyvä eikä huono asia. Se, ovatko tekoälyjärjestelmän tulokset yhteiskunnan kannalta myönteisiä vai kielteisiä, riippuu siitä, miten tekoälyjärjestelmä suunnitellaan, miten ja kuka sitä käyttää ja mihin tarkoituksiin.
- 25 On tietoinen siitä, että ihminen ei pysty tekemään sitä, mitä tekoälyjärjestelmät pystyvät helposti tekemään (esim. tunnistamaan säännönmukaisuuksia valtavista tietomääristä); kun taas tekoälyjärjestelmät eivät pysty tekemään monia asioita, joita ihmiset pystyvät helposti tekemään (esim. ymmärtämään, päättämään, mitä tehdä, ja soveltamaan inhimillisiä arvoja).
- 26 Tunnistaa, että kuvien, kirjoitusten ja musiikin luomiseen suunnitellut tekoälytyökalut ovat riippuvaisia ihmisistä (esim. alkuperäisten parametrien asettaminen ja lopputulosten valitseminen), kun taas ihmiset voivat käyttää tekoälytyökaluja luovutensa lisäämiseksi.

Vuorovaikutus tekoälyjärjestelmien kanssa

Hakusyötteiden muotoilu tekoälysovelluksia käytettäessä on erittäin tärkeä uusi taito, oli sitten kysymyksessä puhetta tunnistava digiavustaja tai ChatGPT:n työkalu. Mitä tarkempi syöte, sen parempi tulos. Käyttäjän on myös hyvä tietää, että jotkut algoritmit on ohjelmoitu tuottamaan käyttäjän mielipidettä tukevia tuloksia. Tällöin vaarana on ns. kaikukammio-ilmio. Siksi käyttäjien tulisi harkita tekoälypohjaisten hakukoneiden käytön etuja ja haittoja.

Tekoälyjärjestelmien ja sovellusten käyttö

Generatiiviset tekoälyjärjestelmät voivat olla erinomainen apu vaikkapa kirjoittajalle. Ne voivat ehdottaa käyttäjälleen erilaisia ideoita ja lähestymistapoja. Käyttäjän on kuitenkin syytä tarkistaa tiedon oikeellisuus ja harkita, noudattaako hän tekoälyn suosituksia vai ei.

Generatiivisten tekoälyn tuottamien vastausten laatua pystytään parantamaan antamalla palautetta, korjaamalla syötettä ja tarkentamalla annettuja tehtäviä ja ohjeita. Tietoverkkoon kytketyt laitteet ja ohjelmistot antavat mahdollisuuden säätää asetuksia esimerkiksi oman yksityisyyden ja tietoturvan parantamiseksi. Internetin sivustoilta latautuu tietokoneelle erilaisia eväs-

teitä (cookies), jotka keräävät ja välittävät monenlaista tietoa eteenpäin. Niiden toimintaa voi säädellä ja niitä voi poistaa.

Erilaisten käänös- ja tulkkausovellusten tarkkuus on parantunut viime vuosien aikana merkittävästi. Niiden tekemät käänökset on kuitenkin syytä tarkistaa väärinkäsitysten välttämiseksi.

Hakukoneiden algoritmit pystyvät alle sekunnissa löytämään usein tuhansittain linkkejä koskien melkein mitä tahansa hakusanaa. Informaation etsijä joutuu puntaroimaan tarkasti, mikä tuhansista löydöistä vastaa informaatiotarvetta, etenkin kun hakutulosten järjestykseen pystytään vaikuttamaan esimerkiksi ostamalla näkyvyyttä tai optimoimalla internetsivusto hakukoneystävälliseksi käyttäen erilaisia teknisiä mahdollisuuksia. Hakutulosten kärkitulokset saattavat heijastaa kaupallisia ja muita intressejä eivätkä ole kyselyn kannalta sopivimpia tuloksia.

Hakukoneita käytettäessä on myös tarpeellista analysoida ja arvioida kriittisesti hakutuloksia, tunnistaa niiden tekijä tai alkuperä, erottaa tosiseikat mielipiteistä ja määrittää, onko löydetty informaatio totuudenmukaista ja vailla taloudellisia, poliittisia tai uskonnollisia intressejä.

Luovissa prosesseissa generatiiviset tekoälytyökalut saattavat olla hyviä avustajia, mutta käyttäjän on syytä säilyttää ohjat omissa käsissään.

Kun kansalainen on vuorovaikutuksessa tekoälyjärjestelmien kanssa, hän...

- 29** Osaa muotoilla hakukyselyitä halutun lopputuloksen saavuttamiseksi vuorovaikutuksessa digiavustajien tai älykkäiden kaiuttimien (esim. Siri, Alexa, Cortana, Google Assistant) kanssa, esim. ymmärtäen, että jotta järjestelmä voi vastata halutulla tavalla, kyselyn on oltava yksiselitteinen ja selkeästi esitetty, jotta järjestelmä voi vastata.
- 30** Kykenee tunnistamaan, että jotkut tekoälyalgoritmit voivat vahvistaa olemassa olevia näkemyksiä digitaalisissa ympäristöissä luomalla ns. kaikukammioita tai informaatiokuplia (esim. jos sosiaalisessa mediassa suositaan tiettyä poliittista ideologiaa, lisäsuositukset voivat vahvistaa kyseistä ideologiaa altistamatta sitä vastakkaisille argumenteille).
- 31** Puntaroi tekoälypohjaisten hakukoneiden käytön hyötyjä ja haittoja (esim. vaikka ne saattavat auttaa käyttäjiä löytämään halutun tiedon, ne voivat vaarantaa yksityisyyden ja henkilötietojen suojan tai altistaa käyttäjän kaupallisille intresseille).

Toimiakseen tekoälyjärjestelmien kanssa luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen...

- 32** On avoin tekoälyjärjestelmille, jotka tukevat ihmisiä tekemään tietoon perustuvia päätöksiä tavoitteidensa mukaisesti (esim. käyttäjät päättävät aktiivisesti, noudattavatko he suositusta vai eivät).
- 33** Pystyy olemaan vuorovaikutuksessa ja antamaan palautetta tekoälyjärjestelmälle (esim. antamalla käyttäjälle arvosanoja, tykkäyksiä, tunnisteita verkkosisällölle), jotta se voi vaikuttaa siihen, mitä se seuraavaksi suosittelee (esim. saadakseen lisää suosituksia samankaltaisista elokuvista, joista käyttäjä on aiemmin pitänyt).
- 35** Osaa muuttaa käyttäjän asetuksia (esim. sovelluksissa, ohjelmistoissa, digitaalisissa alustoissa), jotta tekoälyjärjestelmä voi seurata, kerätä tai analysoida tietoja, estää sen käytön tai hillitää sitä (esim. estää matkapuhelinta seuraamasta käyttäjän sijaintia).
- 36** Tietää, miten ja milloin käyttää konekääntämiskäytäntöjä (esim. Google Translate, DeepL) ja simultaanitulkkaussovelluksia (esim. iTranslate) saadakseen karkean käsityksen asiakirjasta tai keskustelusta. Tietää kuitenkin myös, että kun sisältö vaatii tarkkaa käännöstä (esim. terveydenhuollossa, kaupassa tai diplomaattisissa), saatetaan tarvita tarkempaa käännöstä.

Tekoäly ja tietosuoja

EU:n alueella on voimassa yleinen tietosuoja-asetus (GDPR), jolla asetetaan yrityksille ja organisaatioille henkilötietojen keräämistä, säilytystä ja hallinnointia koskevat tarkat vaatimukset. Henkilötiedoilla tarkoitetaan esimerkiksi henkilön nimeä, osoitetta, ID:n tai passin numeroa, tulotietoja, IP-osoitetta, kulttuurillista profiilia tai terveystietoja. Myös henkilön etninen alkuperä, sukupuolinen suuntautuminen, poliittinen tai uskonnollinen mielipide, ammattiliittoon kuuluminen ja geneettiset ja biometriset tiedot kuuluvat suojauksen piiriin.

Tekoälyjärjestelmiä käytettäessä on hyvä olla tietoinen tietosuoja-asetuksesta ja sen mahdolliseen väärinkäyttöön liittyvistä riskeistä. Biometristen tunnistustekniikoiden käytön vaikutuksia turvallisuuteen on harkittava huolellisesti ennen sormenjälkien tai kasvokuvan rekisteröimistä. Virtuaalisten avustajien ja tekoälyä hyödyntävien laitteiden aktivoimisen mahdolliset seuraukset yksityisyydelle on hyvä arvioida. Tekoälyn käytön myönteiset ja kielteiset vaikutukset on tunnistettava erityisesti henkilötietojen osalta. Julkisesti sosiaalisessa mediassa jaettuja tietoja voidaan käyttää tekoälyjärjestelmien koulutukseen. Esimerkiksi Meta on käyttänyt Instagramiin ja Facebookiin jaettuja kuvia ja tekstejä tekoälyn kouluttamiseen.⁶⁵

Jotta tekoälyjärjestelmien kanssa voi toimia luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen...

- 39** Tietää, että henkilötietojen käsittelyyn sovelletaan paikallisia säännöksiä, kuten EU:n yleistä tietosuoja-asetusta (GDPR). Esimerkiksi ääniyhdydet virtuaaliavustajan kanssa ovat GDPR:n kannalta henkilötietoja ja voivat altistaa käyttäjät tietyille tietosuojan, yksityisyyden ja turvallisuuden liittyville riskeille.
- 40** Punnitsee biometristen tunnistustekniikoiden (esim. sormenjälki, kasvokuvat) käytön hyödyt ja riskit, koska ne voivat vaikuttaa turvallisuuteen tahattomasti. Jos biometriset tiedot vuotavat tai hakkeroidaan, ne vaarantuvat ja voivat johtaa identiteettipetoksiin.
- 41** On tietoinen siitä, että tekoälyjärjestelmät, jotka tukeutuvat käyttäjien henkilötietoihin (esim. ääniavustajat, chatbotit), saattavat kerätä ja käsitellä näitä tietoja enemmän kuin on tarpeen. Tätä voidaan pitää "kohtuuttomana" ja näin ollen se rikkoisi tietosuoja-asetuksessa määritellyä suhteellisuusperiaatetta.
- 42** Punnitsee hyödyt ja riskit ennen virtuaalisen avustajan (esim. Siri, Alexa, Cortana, Google Assistant) tai tekoälypohjaisten esineiden internetin (IoT) laitteiden aktivoimista, koska ne voivat paljastaa henkilökohtaisia päivittäisiä rutiineja ja yksityisiä keskusteluja.

- 43** Punnitsee hyödyt ja riskit ennen kuin antaa kolmansien osapuolten käsitellä henkilökohtaisia tietoja (esim. tiedostaa, että älypuhelimien ääniavustaja, jota käytetään komentojen antamiseen robotti-imurille, voi antaa kolmansille osapuolille –yrityksille, hallituksille, tietoverkkorikollisille – pääsyn tietoihin).
- 44** Tunnistaa sekä myönteiset että kielteiset vaikutukset, joita on kaikkien tietojen ja erityisesti henkilötietojen keräämisellä, koodauksella ja käsittelyllä, tekoälyyn perustuvassa digitaalisessa teknologiassa, kuten sovelluksissa ja verkkopalveluissa.
- 45** On tietoinen siitä, että kaikkea, mitä ihminen jakaa julkisesti verkossa (esim. kuvat, videot, äänet), voidaan käyttää tekoälyjärjestelmien kouluttamiseen. Esimerkiksi kaupalliset ohjelmistoyritykset, jotka kehittävät tekoälyn kasvojen tunnistusjärjestelmiä, voivat käyttää verkossa jaettuja henkilökohtaisia kuvia (esim. perhevalokuvia) kouluttaakseen ja parantaakseen ohjelmiston kykyä tunnistaa automaattisesti kyseiset henkilöt muissa kuvissa, mikä ei ehkä ole toivottavaa (esim. voisi loukata yksityisyyttä).

Tekoälyn sudenkuopat ja eettiset kysymykset

Kansalaisten tulisi olla tietoisia tekoälyn toimintaan liittyvistä riskeistä ja käyttää tekoälysovelluksia kriittisesti. Tekoäly voi välittää kaupallisia tai poliittisia viestejä ja vahvistaa ennakkoluuloja. Generatiiviset tekoälysovellukset on koulutettu valtavalla määrällä materiaalia, joka saattaa sisältää stereotyyppioita, kuten sukupuoli-vinoumia.

Internetin käyttäjien on useimmiten mahdotonta tietää, miten esimerkiksi hakutulosalgoritmit toimivat ja millä perusteella ne esittelevät hakutulokset käyttäjälle.

Aidolta näyttäviä mutta täysin keinotekoisia videoita ja kuvia pystytään luomaan syvävääreennösteknologialla (deepfake).

Euroopan parlamentti on tehnyt jo esityksen ensimmäisestä tekoälysäädöksestä,⁶⁶ ehdotetut asetukset eivät ole vielä voimassa, vaikka kolmikantaneuvotte- luissa on jo löydetty yhteisymmärrys.. Lainsäädännöllä on tarkoitus rajoittaa niitä järjestelmiä, joita saadaan käyttää EU:n sisämarkkinoilla ja tuoda tänne. Tarkoituksena on lisätä kansalaisten luottamusta tuotteisiin eli tekoälyjärjestelmiin EU:n sisällä sekä luoda mahdollisimman toimivat ja vapaat sisämarkkinat.⁶⁷

Jotta tekoälyjärjestelmien kanssa voi toimia luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen...

- 48** Tiedostaa, että tekoälyalgoritmeja ei välttämättä ole muokattu tarjoamaan vain käyttäjän haluamaa tietoa, vaan ne voivat myös sisältää kaupallisen tai poliittisen viestin (esim. kannustaa käyttäjiä pysymään sivustolla, katsomaan tai ostamaan jotain tiettyä asiaa tai jakamaan tiettyjä mielipiteitä). Tällä voi olla myös kielteisiä seurauksia (esim. stereotyyppioiden toistaminen, väärän tiedon jakaminen).
- 49** On tietoinen siitä, että tekoälyn perustana oleviin tietoihin voi sisältyä ennakkoluuloja. Jos näin on, nämä ennakkoluulot voivat automatisoitua ja kärjistyä tekoälyn käytön myötä. Esimerkiksi ammattia koskevat hakutulokset voivat sisältää stereotyyppioita mies- tai naisvaltaisista ammateista (esim. miespuoliset bussinkuljettajat, naispuoliset myyjät).
- 50** On tietoinen siitä, että tekoälyalgoritmit toimivat tavoilla, joita käyttäjät eivät yleensä näe tai ymmärrä helposti. Tätä kutsutaan usein "mustan laatikon" päätöksenteoksi, koska voi olla mahdotonta jäljittää, miten ja miksi algoritmi tekee tiettyjä ehdotuksia tai ennusteita.
- 51** Tietää, että termi deepfakes viittaa tekoälyn tuottamiin kuviin, videoihin tai äänitallenteisiin tapahtumista tai henkilöistä, joita ei ole todellisuudessa tapahtunut (esim. poliitikkojen puheet, julkisten kasvojen pornografisissa kuvissa). Niitä voi olla mahdotonta erottaa aidosta sisällöstä.
- 53** On tietoinen siitä, että EU pyrkii varmistamaan, että tekoälyjärjestelmät ovat luotettavia. Kaikki tekoälyjärjestelmät eivät kuitenkaan ole luotettavia, eikä kaikkia maailmassa kehitettyjä tekoälyjärjestelmiä säännellä EU:n lainsäädännöllä.

55 On tietoinen siitä, että tekoälyjärjestelmät on yleensä kehitetty englanninkielisissä yhteyksissä, mikä tarkoittaa, että ne saattavat toimia epätarkemmin muissa kuin englanninkielisissä yhteyksissä. Esimerkiksi tekoälypohjaiset automaattiset käännösjärjestelmät toimivat paremmin usein käytettyjen kielten kanssa (esim. englannista espanjaan) kuin harvemmin käytettyjen kielten kanssa (esim. sloveenista suomeen).

56 Tiedostaa, että tekoälyjärjestelmiä kehittävät tyypillisesti tietynlaiseen väestöosaan kuuluvat henkilöt (esim. valkoiset miehet, jotka kuuluvat korkeamman tulotason maiden ylempiin sosioekonomisiin ryhmiin), mikä voi merkitä sitä, että heidän kehittämissään järjestelmissä ei oteta yhtä hyvin huomioon naisten, eri etnisiin vähemmistöryhmiin kuuluvien henkilöiden, alempien sosioekonomisten ryhmien edustajien, digitaalista esteettömyyttä vaativien henkilöiden (esim. vammaisten, toimintarajoitteisten henkilöiden) tai alhaisemman tulotason maiden kansalaisten tarpeita.

Tekoäly ja etiikka

Käyttäjien on syytä tarkastella tekoälypalveluja myös eettisestä näkökulmasta ja huomioida niiden mahdolliset ympäristö- ja yhteiskunnalliset vaikutukset. On tilanteita, joissa pelkästään tekoälypohjainen päätöksenteko on kyseenalaista ja joissa ihmisen väliintulo on välttämätöntä, etenkin jos ne ovat tärkeitä henkilön hyvinvoinnille tai tulevaisuudelle (esim. oppilasarviointi).

Uusi teknologia pyrkii usein helpottamaan ja automatisoimaan työtä, millä on vaikutuksia työpaikkoihin.

Tekoälysovellusten myötä voidaan jotkut yksinkertaiset ja toistuvat työtehtävät siirtää esimerkiksi robottien tai asiakaspalveluchatbottien tehtäväksi. Toisaalta tekoälyn kehitystarpeet luovat uusia työmahdollisuuksia.

Eettisten periaatteiden tulisi olla olennainen osa tekoälyjärjestelmien kehittämistä ja käyttöönottoa. Eettisten periaatteiden tulisi edistää ihmisen toimijuutta. Lisäksi tekoälyjärjestelmien avoimuutta, syrjimättömyyttä, esteettömyyttä, ennakkoluulottomuutta ja oikeudenmukaisuutta tulisi nykyisestä edistää.

Jotta tekoälyjärjestelmien kanssa voi toimia luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen...

57 Pohtii tekoälyjärjestelmien eettisiä seurauksia koko niiden elinkaaren ajan. Näitä ovat ympäristövaikutukset (digitaalisten laitteiden ja palvelujen tuotannon ympäristövaikutukset) ja yhteiskunnalliset vaikutukset, kuten työelämän verkottuminen, joka näkyy digitaalisten alustojen lisääntyvänä käyttönä työntekijöiden ja asiakkaiden välisessä yhteydenpidossa. Yhteiskunnallisiin vaikutuksiin kuuluvat myös algoritminen hallinta, joka voi rajoittaa työntekijöiden yksityisyyttä tai oikeuksia sekä halpatyövoiman käyttö kuvien merkitsemiseen tekoälyjärjestelmien kouluttamiseksi.

58 On valmis pohtimaan tekoälyjärjestelmiin liittyviä eettisiä kysymyksiä (esim. missä yhteyksissä, kuten rikollisten tuomitsemisessa, tekoälysuosituksia ei pitäisi käyttää ilman ihmisen väliintuloa).

60 On tietoinen siitä, että tekoälyyn perustuvilla teknologioilla voidaan korvata joitakin ihmisen tehtäviä (esim. asiakaspalvelu), mikä saattaa johtaa työpaikkojen menetyksiin tai uudelleenjärjestelyihin, mutta että uusia työpaikkoja saattaa syntyä uusien tarpeiden tyydyttämiseksi.

61 Ottaa eettisyyden yhdeksi peruspilariksi kehitettäessä tai otettaessa käyttöön tekoälyjärjestelmiä (mukaan lukien ihmisen toimijuus ja päätösvalta, avoimuus, syrjimättömyys, esteettömyys sekä ennakkoluulot ja oikeudenmukaisuus).

Ihmisen toimijuutta ja päätösvaltaa koskevat asenteet

Generatiiviset tekoälysovellukset kehittyvät nopeasti, ja uudet päivitykset ja sovellukset toimivat aikaisempia paremmin, monipuolisemmin ja tehokkaammin. Niiden kehitystä on hyvä seurata, niitä kannattaa kokeilla ja testata. Samalla on syytä pohtia niiden tuomaa hyötyä sekä käyttöön liittyviä riskejä ja haasteita. Tekoälyn tuottamiin vastauksiin, suosituksiin ja tuloksiin kannattaa suhtautua kriittisesti, ja käyttäjän pitäisi säilyttää pää-

tösvalta siitä, noudattaako niiden antamia suosituksia vai ei.

Monet yritykset ja viranomaiset ovat automatisoineet tekoälyn avulla asiakasta koskevaa päätöksentekoa. Siksi on hyvä tietää, että EU:n kansalaisilla on oikeus olla joutumatta pelkästään algoritmeihin perustuvan päätöksen kohteeksi. Esimerkiksi pankki saattaa käyttää tekoälyä lainatarjouksen hyväksymisessä. EU:n kansalaisella on kuitenkin oikeus pyytää, että pankin henkilökunnan jäsen tarkistaisi automatisoidun päätöksen.

Jotta tekoälyjärjestelmien kanssa voi toimia luottavaisesti, kriittisesti ja turvallisesti, esimerkkeinä mainittakoon, että kansalainen...

- 62** Suhtautuu avoimesti tekoälyjärjestelmiin, jotka auttavat ihmisiä tekemään tietoon perustuvia päätöksiä tavoitteidensa mukaisesti (esim. käyttäjät päättävät aktiivisesti, noudattavatko he suositusta vai eivät).
- 63** Tiedostaa, että vaikka tekoälyjärjestelmien soveltaminen monilla aloilla ei yleensä ole kiistanalaista (esim. tekoäly, joka auttaa torjumaan ilmastomuutosta), tekoäly, joka on suoraan vuorovaikutuksessa ihmisten kanssa ja tekee heidän elämäänsä koskevia päätöksiä, voi usein olla kiistanalaista (esim. ansioluetteloiden lajitteluohjelmistot rekrytointimenettelyjä varten, kokeiden pisteytys, joka voi määrittää pääsyn koulutukseen).
- 64** Tietää, että kaikilla EU:n kansalaisilla on oikeus olla joutumatta täysin automatisoidun päätöksenteon kohteeksi (esim. jos automaattinen järjestelmä hylkää luottohakemuksen, asiakkaalla on oikeus pyytää, että henkilö tarkistaa päätöksen).⁶⁶
- 65** Punnitsee hyötyjä, joita saadaan, mikäli tekoälyjärjestelmiä käytetään parantamaan inhimillisen vuorovaikutuksen laatua viestinnässä (esim. tekoälyn tuottamien vastausten käyttö sähköposteihin saattaa vaarantaa vuorovaikutuksen epäinhimillistymisen).
- 70** On tietoinen siitä, että joskus paras tapa hallita tekoälyjärjestelmää (esim. itsensä ja muiden suojelemissi) on olla olematta vuorovaikutuksessa sen kanssa tai kytkeä se pois päältä.
- 71** On kiinnostunut kokeilemaan erityyppisiä tekoälyjärjestelmiä omien henkilökohtaisten tarpeidensa mukaan (esim. virtuaalinen avustaja, kuva-analyysohjelmistot, puhe- ja kasvojen tunnistusjärjestelmät, autonomiset autot, "ruumiillistettu" tekoäly kuten robotit).
- 72** Haluaa jatkuvasti oppia, kouluttaa itseään ja pysyä ajan tasalla tekoälystä (esim. ymmärtää, miten tekoälyalgoritmit toimivat; ymmärtää, miten automaattinen päätöksenteko voi olla puolueellista; erottaa toisistaan realistisen ja epärealistisen tekoälyn; ja ymmärtää eron tekoälyn suppean älykkyyden ja yleisen älykkyyden välillä. Edellinen tarkoittaa nykypäivän tekoälyä, joka kykenee suppeisiin tehtäviin, kuten pelien pelaamiseen, jälkimmäinen ihmisälyä parempaa tekoälyä, joka on edelleen tieteellistä fiktiota).
- 73** On avoin ja utelias nykypäivän kehittyviä teknologioita ja sovelluksia kohtaan (esim. lukee arvosteluja virtuaalitodellisuudesta, pelaamisesta ja tekoälystä) ja keskustelee tarkoituksella niiden käytöstä muiden ihmisten kanssa.

Digi- ja väestötietovirasto on julkaissut oppaan Vastuullisen tekoälyn hyödyntämisen.⁶⁹ Oppaaseen on liitetty kattavia muistilistoja organisaatioiden tekoälyn käytön tueksi.

Generatiivista tekoälyä hyödyntävien yritysten ja yhteisöjen olisi hyvä sopia yhteisistä pelisäännöistä tekoälyn käytössä. Henkilöstölle tulisi antaa selkeät ohjeet siitä, mihin tekoälyä voi käyttää, ja mitä pitäisi välttää.

Esimerkki tekoälyohjeistuksesta henkilöstölle

- Älä koskaan jaa henkilötietoja tai mitään muita luottamuksellisia tietoja GenAI-sovelluksen kanssa.
- Arvioi aina kriittisesti kaikkia verkossa saatavilla olevia GenAI:n tuottamia vastauksia mahdollisten ennakkoluulojen ja asiasisällöltään virheellisen informaation varalta.
- Älä koskaan kopioi suoraan GenAI:n tuotoksia julkisiin asiakirjoihin tai sähköposteihin ilman tarkistusta.
- Älä koskaan luota verkossa saatavilla oleviin GenAI-malleihin kriittisissä ja kiireellisissä prosesseissa.

⁵⁵ Vuorikari, R., Kluzer, S. and Punie, Y., DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-48883-5, doi:10.2760/490274, JRC128415.

⁵⁶ Euroopan neuvosto (2018) Elinikäisen oppimisen avaintaidot: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_2018.18901.0001.01.ENG&toc=OJ.C:2018:189:TOC

⁵⁷ Vuorikari, R., Kluzer, S. and Punie, Y., DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-48883-5, doi:10.2760/490274, JRC128415.

⁵⁸ UNESCO (2023) Guidance for generative AI in education and research. <https://unesdoc.unesco.org/ark:/48223/ptf0000386693>

⁵⁹ OpenAI site (2023) <https://openai.com>

⁶⁰ Clario (ladattu 19.12.2023) Which data can companies actually collect? <https://clario.co/blog/which-company-uses-most-data/>

⁶¹ UNICEF (Ladattu 19.12.2023). Policy guidance on AI for children. <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>

⁶² Habli, Ibrahim, Lawton, Tom & Porter, Zoe. (2020). Artificial intelligence in health care: accountability and safety. Bulletin of the World Health Organization, 98 (4) , 251 - 256. World Health Organization. <http://dx.doi.org/10.2471/BLT.19.237487>

⁶³ Konam, S. (2022) Where did IBM go wrong with Watson Health?. Quartz. <https://qz.com/2129025/where-did-ibm-go-wrong-with-watson-health>

⁶⁴ Alamäki, A. (2023). Tekoälyllä on merkittävä potentiaali terveyssektorilla. eSignals. <https://esignals.fi/kategoria/tutkimus-ja-kehittaminen/tekoalylla-on-merkittava-potentiaali-terveyssektorilla/#calbee49>

⁶⁵ Fowler, G. (2023) Your instagrams are training AI. There's little you can do about it. The Washington Post. <https://www.washingtonpost.com/technology/2023/09/08/gmail-instagram-facebook-trains-ai/>

⁶⁶ Euroopan parlamentti (2023). EU:n tekoälysäädös on ensimmäinen laatuaan. <https://www.europarl.europa.eu/news/fi/headlines/society/20230601STO93804/eu-n-tekoalysaad-os-on-ensimmainen-laatu-aan>

⁶⁷ Helsingin yliopisto (2021) Tarvitseeko tekoäly kansainvälisiä säätäjiä? <https://www.helsinki.fi/fi/uutiset/tekoaly/tarvitseeko-tekoaly-kansainvalisia-saantajia>

⁶⁸ European Commission (ladattu 19.12.2023) Are there restrictions on the use of automated decision-making? https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/dealing-citizens/are-there-restrictions-use-automated-decision-making_en

⁶⁹ Suomi.fi (ladattu 19.12.2023) Tekoälyn vastuullinen hyödyntäminen. <https://www.suomi.fi/oppaat/vastuullinen-tekoaly/muistilista>

6. Nuoret, oppiminen, sosiaalinen media & tekoäly

KARI KIVINEN

Tammikuussa 2007 Steve Jobs esitteli laitteen, joka oli iPodin, matkapuhelimen ja internet-selaimen yhdistelmä. IPhonen myötä älypuhelimet ovat mullistaneet meidän jokapäiväisen arkemme. Lähes kaikilla suomalaisilla on nykyään tehokas mobiililaitte käytössään. Juuri tällä hetkellä on meneillään samanlainen mullistus OpenAI:n julkaistua kaikkien saataville ilmaisen generatiivisen tekoälysovelluksen ChatGPT:n.

Ebrandin kyselyn mukaan suomalainen keski-ikäinen nuori käyttää internetiä 35–42 tuntia viikossa, joista sosiaalisen median palveluita noin 15–20 tuntia viikossa.⁷⁰ Tämä tarkoittaa sitä, että suuri osa nuorista saattaa olla viikoittain enemmän netissä kuin esimerkiksi koulussa. Saman kyselyn mukaan suomalaisnuorten suosituimmat sosiaalisen median palvelut ovat WhatsApp, YouTube, Instagram, Spotify, Snapchat, TikTok, Facebook, Discord, Pinterest, Jodel ja Steam.

Sosiaalisen median käyttämät suositelualgoritmit perustuvat tekoälyn tehokkaaseen käyttöön. Samalla kun käyttäjä hakee tai jakaa mitä tahansa sisältöä, niin hänestä jää digitaalinen jalanjälki, jolla täydennetään alustojen käyttäjäprofiilia, jota taas käytetään esimerkiksi mainosten kohdentamiseen ja generatiivisen tekoälyn kouluttamiseen.

Eniten viime vuosina suosiota on lisännyt TikTakin käyttö. Vaikka monet viralliset tahot ovat kieltäneet TikTakin käytön työtietokoneissa, sen suosio erityisesti nuorten parissa on kasvanut joka puolella maailmaa. TikTokilla on jo yli miljardi käyttäjää, ja se on alkanut mainostaa itseään ”mikro-oppimisalustana”. Videoita katsomalla miljoonat käyttäjät oppivat uusia tietoja ja taitoja päivittäin.

Sama pätee erityisesti YouTubeen. Alustoilta on



helppo etsiä mitä tahansa opetusvideoita melkein mistä tahansa aiheesta. Joukkoon mahtuu myös paljon virheellistä tai vääristynyttä informaatiota. Toisaalta tätä ”video-oppimista” voisi hyödyntää tehokkaammin myös kouluopetuksessa.

Digilaitteiden käytön räjähdysmäisen kasvun seurauksena kirjojen ja printtimedian lukemiseen käytetty aika on supistunut⁷¹ ja kiinnostus esimerkiksi pidempien tekstien ja tietokirjojen lukemiseen on laskenut. Tiedonhaku on siirtynyt internetiin. Vaikka monet nuoret ovat taitavia laitteiden käyttäjiä, niin netti-informaation kriittinen lukutaito ja informaation analysointi on yllättävän monelle vaikeaa. Critical-ryhmän tutkimusten mukaan⁷²

- noin 40 prosenttia kuudesluokkalaisista ei kyseenalaistanut kaupallisen tekstin kirjoittajan asiantuntijuutta lainkaan;
- suurin osa tutkimukseen osallistuneista nuorista ei osaa tulkita harhaanjohtavia diagrammeja;
- lähes 30 prosentilla lukiolaisista oli vaikeuksia arvioida nettitekstien luotettavuutta.

Critical-ryhmän suosituksen mukaan jokainen nuori tarvitsee kriittistä lukutaitoa eli taitoa hakea tietoa eri medioista, arvioida tekstien luotettavuutta ja vertailla tekstien sisältämää informaatiota. Kriittistä lukutaitoa ja peruslukutaitoa on tuettava rinnakkain. Lukemisen haasteisiin tulee tarjota tukea mahdollisimman varhaisessa vaiheessa.

Nuorille suunnatut sosiaalisen median alustat tarjoavat monenlaisia tekoälypohjaisia sovelluksia. Esimerkiksi Snapchatissa on jokaiselle käyttäjälle oma henkilökohtainen My-AI-ystävä, jonka kanssa käyttäjät voivat jutella, kysellä asioita ja jolta voi saada suosituksia. Toisaalta tutkijat varoittavat tekoälyn liiallisesta käytämisestä vuorovaikutustilanteiden automatisoimiseen, sillä se riistää ihmisten välisen kanssakäymisen psyykkiset ja sosiaaliset hyödyt.⁷³ Pandemia on jo antanut meille välähdyksen siitä, mitä tapahtuu, kun nämä siteet katkeavat, kuten Survey Center on American Life -keskuksen johtaja Daniel Coxin on todennut. Kun vietämme vähemmän aikaa toistemme kanssa, menetämme tottumuksen tulla toimeen keskenämme yhteisissä tiloissa. Siksi tekoäly on niin huono korvike todelliselle vuorovaikutukselle. Meidän olisi vietettävä enemmän aikaa toistemme kanssa.

Algoritimpohjaiset sovellukset, kuten esimerkiksi Internetin hakukoneet, ovat osa jokaisen nuoren ja myös jokaisen opettajan arkipäivää. Siksi opettajien tulisi ottaa tämä huomioon omassa työssään ja opetuksessaan. Ebranding-kyselyn mukaan peräti 79 prosenttia nuorista vastaajista oli sitä mieltä, että sosiaalisen median käyttöä tulisi opettaa nuorille kotona tai koulussa tai oppilaitoksessa.⁷⁴

Unescon raportin mukaan GenAI-työkalut muuttavat tapaa, jolla opetus- ja oppimissisältöä voidaan tuottaa ja tarjota.⁷⁵ Tulevaisuudessa ihmisen ja tekoälyn välisen keskustelujen avulla tuotetusta sisällöstä voi tulla yksi tärkeimmistä tiedontuotannon lähteistä. Tämä todennäköisesti heikentää entisestään oppijoiden suoraa sitoutumista opetussisältöön, joka perustuu ihmisten luomiin ja validoihin resursseihin, oppikirjoihin ja opetussuunnitelmiin.

Oppimisessa luotettavan tiedon etsimisen prosessi on tärkeämpi kuin huoliteltu lopputulos; mitkä ovat keskeiset informaation lähteet, miten tieto on rakentunut, mihin evidenssiin se perustuu ja mitä johtopäätöksiä siitä voi tehdä?

Opettajien onkin pohdittava uudesta näkökulmasta, minkälaisia tehtäviä opiskelijoille kannattaa antaa. Kuka tahansa kykenee luomaan GenAI:n avulla upean esseen mistä tahansa aiheesta ilman minkäänlaista ymmärrystä asian sisällöstä. Miten antaa opiskelijoille mahdollisuus käyttää GenAI:n potentiaalia, mutta samalla varmistaa, että opiskelijat ovat itsenäisesti löytäneet informaation lähteet ja tarkistaneet esitettyjen tietojen oikeellisuuden?

Stanfordin yliopiston valtio-opin professori Rob Reich on esittänyt provosoivan kysymyksen: "Onko ge-

neratiivinen tekoäly verrattavissa laskimeen luokkahuoneessa, vai onko se haitallisempi työkalu?"⁷⁶

Oppilaat ratkaisevat nykyään monimutkaisia laskutoimituksia ja graafisia yhtälöitä laskimien avulla. Se ei kuitenkaan ole johtanut matemaattisen peruslaskutoimituksen poistamiseen opetussuunnitelmasta: Oppilaat osaavat edelleen laskea jakolaskuja tai ratkaista yhtälöitä myös ilman laskinten apua. Toisaalta kirjoittaminen on Reichin mukaan tapa oppia ajattelemaan ja hän pohtiikin, voiko kirjoittamisen ulkoistaminen tekoälylle vahingoittaa oppilaiden kriittisen ajattelun kehittymistä? (tästä enemmän artikkelissa Matt Croninin artikkelissa "Kehitkö tekoälyn neuroverkkoa, vai omaasi?")

Opetushallitus on linjannut artikkelissaan:⁷⁷

"Tekoälyä käytetään jatkossa merkittävästi eri työtehtävissä. Se muuttaa työnkuvia, tehostaa työtä ja tuottaa kokonaan uudenlaisia ratkaisuja. Tämä ei kuitenkaan korvaa tarvetta perustaitojen sekä oppimaan oppimisen taitojen opetukselle. Tekoälysovellusten laajentuminen haastaa erityisesti kriittistä lukutaitoa, mikä puolestaan perustuu yleissivistykseen. Tekoälyyn tutustuminen edistää teknologista osaamista ja tukee monipuolisen lukutaidon kehittymistä. Tekoälyn periaatteiden pohtiminen on tärkeää, jotta oppija voi tunnistaa ja arvioida, mikä on mahdollisesti tekoälyn tuottamaa informaatiota tai disinformaatiota."

Tutkija Ilkka Tuomi on julkaissut useita koulutuksen tulevaisuutta ja teknologian merkitystä pohtivia julkaisuja.⁷⁸ Mielipidekirjoituksessaan hän toteaa:⁷⁹

"Sisältöjen oppimisen sijaan tekoälyn aikakaudella koulutuksen tavoitteena on oltava ihmisen toimijuuden vahvistaminen. Se tarkoittaa erityisesti kykyä asettaa ja toteuttaa ihmisen tärkeinä ja arvokkaina pitämiä tavoitteita. Tulevaisuuden osaaminen on teknologian tukeman yksilön osaamista ja yhteisöllisen aikaansaamisen kykyä. Osaaminen ei sijaitse oppijan päässä, eikä sitä sieltä kannata etsiä. Tekoälyä ei voi jättää kännykkäparkkiin, sillä siitä on tulossa erottamaton osa oppijaa ja ajattelun prosesseja."

Useimmissa tekoälyä koskevissa politiikoissa, strategioissa ja suuntaviivoissa lapset ja nuoret mainitaan vain ohimennen. Siksi UNICEF julkaisi 2021 toimintaohjeet edistääkseen lasten oikeuksia. Tavoitteena oli lisätä tekoälypolitiikkaa laativien tahojen tietoisuutta lasten oikeuksista ja siitä, miten tekoälyjärjestelmät voivat tukea tai heikentää näitä oikeuksia. Toimintaohjeissa tarkastellaan tekoälyjärjestelmien kehittymistä ja pohditaan, miten ne vaikuttavat lapsiin.⁸⁰

Ohjeissa esitetään lapsen oikeuksia koskevaan yleisopimukseen nojautuen yhdeksän vaatimusta lapsilähtöiselle tekoälylle:

1. Tuetaan lasten kehitystä ja hyvinvointia.
2. Varmistetaan lasten osallisuus ja osallistaminen.
3. Asetetaan etusijalle oikeudenmukaisuus ja syrjimättömyys lapsia kohtaan.

- Suojataan lasten tietoja ja yksityisyyttä.
- Varmistetaan lasten turvallisuus.
- Tarjotaan avoimuutta, selitettävyyttä ja vastuuvollisuutta lapsiin nähden.
- Annetaan hallituksille ja yrityksille tietoa tekoälystä ja lasten oikeuksista.
- Valmistetaan lapsia tekoälyn nykyiseen ja tule-

vaan kehitykseen.

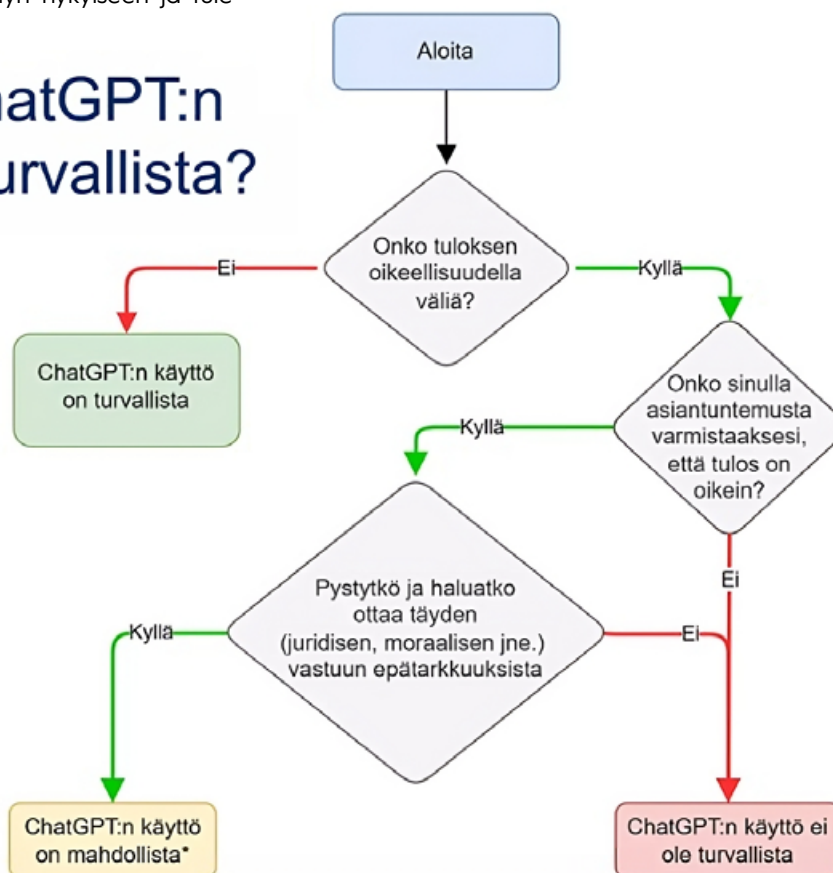
- Luodaan kannustava ympäristö.

Unescon opiskelijoille suunnatussa julkaisussa "ChatGPT and AI in Higher Education Quick Start guide"⁶¹ esitellään oiva päätöksentekopuu ChatGPT:n käytöstä. Sen on suomeksi kääntänyt Matleena Laakso.⁶²

Milloin ChatGPT:n käyttö on turvallista?

Flowchart devised by Aleksandr Tiulkanov, AI and Data Policy Lawyer, 2023, CC BY, suom. energia.vamk.fi www.icsaic.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

* Muista tarkistaa jokaisen syötetyn sanan ja lauseen tarkkuus ja käytä tervettä järkeä.



Eri oppilaitokset ovat alkaneet julkaista ohjeita ja suosituksia opiskelijoilleen tekoälyn käytöstä. Tässä esimerkkinä seuraavat taulukot, jotka on muokattu Gradian⁸³

suosituksista opiskelijoille tekoälyn hyödyntämiseen opinnoissa.

H Y Ö D Y N N Ä	Vuorovaikutteisuus	Valjasta tekoäly esim. testaamaan tietojasi. Pyydä tekoälyä luomaan sinulle tasosi mukaisia kertaustehtäviä, pulmia tai monivalintakysymyksiä tekstistä, jota parhaillaan luet.
	Sparrailu	Tekoäly on väsymätön sparraaja. Voit kysyä siltä kysymyksiä (tarkista kuitenkin vastaukset) tai vinkkejä ihan arkisistakin asioista (hyvinvointi, opiskelutekniikat, jne.)
	Ideointi	Tekoälyn avulla voi ideoida erilaisia kuvituksia ja tehtäviä melkein mistä aiheesta hyvänsä. Voit kysyä vinkkejä eri aiheisiin ja pallotella sen kanssa ideoita.
	Oppiminen	Oppiminen on usein pohtimista ja oivaltamista. Jos on oppinut jonkun asian, niin sen pystyy toistamaan ja opittua voi soveltaa jatkossa myös muihin tilanteisiin. Tekoälyn avulla saa nopeasti valmiita lopputuloksia – muttei välttämättä opi itse yhtään mitään!
	Syötetaito	Mitä paremmin muotoiltu syöte – sen parempi lopputulos. Ei kannata tyytyä ensimmäiseen tuotokseen, vaan hioa sitä paremmaksi syötettä jalostamalla.

Tekoälysovelluksia käytettäessä opiskelun tukena on hyvä huomioida ainakin seuraavia seikkoja:

I O M O J U H	Tietoturva	Älä syötä sovelluksiin mitään henkilökohtaista tietoa itsestäsi tai muista. Sovellukset tallettavat tiedot ja saattavat käyttää niitä muissa yhteyksissä
	Tekijänoikeudet	Sovelluksiin ei saa syöttää toisten luomaa materiaalia ilman lupaa. Et voi esittää tekoälyn tuottamaa materiaalia omanasi.
	Vastuullisuus	Huomioi opinahjosi eettiset ohjeet! Tarkista aina huolellisesti tekoälyn tuottama sisältö. Sen jakaminen on aina sovelluksen käyttäjän vastuulla
	Virheellinen sisältö	Tekoälyn tuottama sisältö ei välttämättä pidä paikkansa. Siksi kannattaa suhtautua tekoälyn tuottamaan aineistoon rakentavan kriittisesti ja tarkistaa esim. kaikki lähdetiedot.
	Oppiminen	Oppiminen on tärkeintä – pelkästään kopioimalla ei opita tulevaisuudessa tarvittavia taitoja.

Lisäoppia tekoälystä

Santeri Kallio on julkaissut erittäin selkeän ja havainnollisen ”Mitä on generatiivinen tekoäly – GenAI Opas”-verkkokurssin ja oppaan generatiivisen tekoälyn perusteista.⁸⁴

Jos haluaa oppia ymmärtämään tekoälyä vähän syvemmin, niin kannattaa suorittaa MinnaLearnin ja Helsingin yliopiston ilmainen verkkokurssi ”The Elements of AI”.⁸⁵ Kurssin tavoitteena on rohkaista jokaista ikään

ja taustaan katsomatta oppimaan tekoälyn perusteet: mitä tekoäly on, mitä sillä voi (ja ei voi) tehdä ja miten luoda tekoälymenetelmiä. Kurssit koostuvat teoriaosuuksista ja käytännön harjoituksista. Kurssi on käännetty monelle kielelle, ja sen on suorittanut jo yli miljoona tekoälystä kiinnostunutta oppijaa.

Verkosta löytyy monenlaisia AI-oppaita ja sanastoja esim. Time lehden A to Z-artikkeli⁸⁶. Erityisesti generatiivisen tekoälyn kehitysvauhti on niin kiivas, että kannattaa etsiä mahdollisimman tuore ja luotettava lähde.

⁷⁰ SoMe ja Nuoret (2022) – katsaus nuorten sosiaalisen median käytöstä. Ebrand. <https://wordpress.ebrand.fi/somejanuoret2022/tiivistelma/>

⁷¹ Lukukeskus – 5 faktaa lukemisesta: <https://lukukeskus.fi/5-faktaa-lukemisesta/>

⁷² Critical-ryhmän politiikkasuositus: https://educritical.fi/wp-content/uploads/2023/03/politiikkasuositus_CRITICAL_27_03_2023.pdf

⁷³ Sheidlower, N. (2023) Nearly a third Gen Z said technology is making them lonelier – and it’s making many spend more. <https://www.businessinsider.com/gen-z-loneliness-technology-ai-social-media-spending-inflation-millennials-2023-12>

⁷⁴ SoMe ja Nuoret 2022 – katsaus nuorten sosiaalisen median käytöstä. Ebrand. <https://wordpress.ebrand.fi/somejanuoret2022/tiivistelma/>

⁷⁵ UNESCO (2023) Guidance for generative AI in education and research. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

⁷⁶ Chen, C. (2023) AI will transform teaching and learning. Let’s do it right. Stanford Accelerator for Learning. <https://acceleratelearning.stanford.edu/story/ai-will-transform-teaching-and-learning-lets-get-it-right/>

⁷⁷ Opetushallitus (ladattu 18.12.2023) Tekoälyä hyödyntäviin sovelluksiin tutustuminen koulussa ja oppilaitoksissa. <https://www.oph.fi/fi/opettajat-ja-kasvattajat/joko-teidan-tyoyhteisossanne-keskustellaan-tekoalysta-0>

⁷⁸ Tuomi, I., Cachia, R. and Villar Onrubia, D., On the Futures of Technology in Education: Emerging Trends and Policy Implications, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/079734, JRC134308.

⁷⁹ Tuomi, I. (2023) Tekoälyä ei voi jättää koulun kännykkäparkkiin. HS. <https://www.hs.fi/mielipide/art-2000009622272.html>

⁸⁰ UNICEF (2021) Policy guidance on AI for children. <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>

⁸¹ UNESCO (2023) ChatGPT and Artificial Intelligence in higher education. Quick start guide. https://www.iesal.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

⁸² Laakso, M. (ladattu 19.12.2023) Tekoäly opetuksessa. <https://www.slideshare.net/MatleenaLaakso/tekoaly-opetuksessa-6923>

⁸³ Gradia (2023) <https://www.gradia.fi/opiskelijalle/uudelle-opiskelijalle>

⁸⁴ Kallio, S. (2023) Mitä on generatiivinen tekoäly – GenAI Opas. <https://santerikallio.com/genai-opas/>

⁸⁵ Elements of AI (ladattu 19.12.2023) <https://course.elementsofai.com/fi/>

⁸⁶ Perrigo, B. (2023) A to Z of Artificial Intelligence. Time. <https://time.com/6271657/a-to-z-of-artificial-intelligence/>



7. Tekoölyn ongelmat

KARI KIVINEN

Tekijänoikeudet

Generatiiviset tekoölymallit on koulutettu erittäin suuresta määrästä tekstiä, koodia, kuvia jne. Ne on usein poimittu Internetistä ja yleensä ilman omistajan lupaa. Itse asiassa emme tiedä tarkasti, mitä materiaalia koulutukseen on käytetty.

GenAI-yrityksiä vastaan on jo nostettu useita kanteita, joissa on riitautettu tekijänoikeussuojattujen teosten käyttö koulutusmateriaalina. Samanaikaisesti kansallisella ja kansainvälisellä tasolla harkitaan sääntely⁶⁸- ja itse-sääntelyaloitteita, joilla pyritään vastaamaan generatiivisen tekoölyn tuomiin haasteisiin ja mahdollisuuksiin.

EU:n ensimmäinen tekoölyä koskeva sääntelykehys, tekoölylaki⁶⁹, on hyväksytty Euroopan neuvostossa joulukuussa 2023. EU:ssa ollaan valmistelemaan tekoölyalan yrityksille avoimuusvaatimuksia, joiden mukaan niiden on toimitettava yksityiskohtaiset asiakirjat koulutustarkoituksiin käytetystä tekijänoikeussuojatusta sisällöstä.

Generatiivisen tekoölyn avulla luodut tekstit ja kuvat saattavat loukata jonkun tekijänoikeuksia. Kansainvälinen käytäntö kuitenkin on vielä tällä hetkellä, tekijänoi-

keuksilla voidaan suojata hetkellä suojata vain ihmisen tekemiä tuotoksia. Tekoölyn tekemät tuotokset eivät siis nauti tekijänoikeussuojaa. Meneillään on kuitenkin useita kansainvälisiä oikeustapauksia, jotka saattavat muuttaa nykyistä käytäntöä. Jotkut yksittäistapaukset saattavat noudattaa perinteisiä tekijänoikeussääntöjä. Esimerkiksi, jos tiedetään, että tekoölyjärjestelmä on koulutettu alkuperäisen teoksen avulla, ja tekoölytuotos on identtinen alkuperäisen teoksen kanssa, niin on hyvin mahdollista, että tekoölytuotos loukkaa teoksen tekijänoikeussuojaa.

INPUT - Generatiivisen tekoölyn kouluttamiseen liittyvät kysymykset

GenAI-yritykset kouluttavat järjestelmiään tietokokonaisuuksien pohjalta, joihin voi kuulua mm. tekoölybottien verkosta keräämää sisältöä. Tämä herättää useita tekijänoikeuskysymyksiä, kuten esimerkiksi

- GenAI-yritysten tekoölyn koulutukseen käytetyn sisällön avoimuus ja sen dokumentointi
- Ratkaisut, joilla estetään tekijänoikeuden suoja-

man sisällön keräämisen verkosta (scraping). Verkkosisällön osalta useat GenAI-yritykset ovat jo ilmoittaneet noudattavansa ns. robotti.txt-protokollaa⁷⁰ ja siihen liittyviä rajoituksia, jotka koskevat verkkosivujen indeksointia ja aineiston keräämistä. Sen lisäksi muita teknisiä ja oikeudellisia lähestymistapoja on kehitteillä tai harkinnassa.

- Mahdollisuus lisensoida laillisesti sisältöä GenAI-koulutusmateriaalin esimerkiksi kollektiivisen oikeudenhallinnan järjestelmien kautta.
- Kehittyvä oikeuskäytäntö: oikeudenhaltijat ovat nostaneet useita kanteita GenAI-yri tyksiä vastaan, jotka ovat jo käyttäneet heidän sisältöään koulutusmateriaalina. Voitaisiko tällainen käyttö sisältää tekijänoikeutta koskeviin poikkeuksiin?
- Kansainvälinen ulottuvuus, joka liittyy tekijänoikeuksien alueellisuuteen. Osa EU:ssa palveluja tarjoavat GenAI-yritykset toimivat sellaisilla lainkäyttöalueilla, jotka eivät välttämättä rajoita tekijänoikeuksien suojaaman aineiston käyttöä koulutustietona.
- Syötetietojen käyttö tuotoksen hyväksyttävyyden todisteena. Syötteessä voisi rajata esimerkiksi, että tuotoksen aineisto on laillisesti saatavilla ja käytössä ja että se ei ole puolueellista, ennakkoluuloista tai epätasapainoista.

OUTPUT - Generatiivisen tekoälyn tuotokset

GenAI-palveluiden tuotoksia voidaan käyttää avustamaan ihmistä hänen innovatiivisissa ja luovissa tehtävissä. Tämä herättää väistämättä erilaisia tekijänoikeuksiin liittyviä kysymyksiä, kuten esimerkiksi:

GenAI:n tuotosten tekijänoikeussuoja: Muutamia teollis- ja tekijänoikeusvirastot ovat antaneet suunta- viivoja kriteereistä, joiden perusteella tekoälyn tuottamat tuotokset voitaisiin suojata tekijänoikeudella. Tällä hetkellä ainoastaan ihmisten tekemä luova työ nauttii tekijänsuojaa.

Tekijänoikeussuoja GenAI:n tuotoksia vastaan: Jotkin oikeudenhaltijat pohtivat myös, miten heitä voidaan suojella siltä, että he eivät joudu kilpailemaan heidän teostensa pohjalta tuotettujen tai niitä jäljittelevien tuotosten kanssa.

GenAI-työkalujen käyttö tekijänoikeuksia loukkaaviin toimiin ja rikoksiin ja niiden estäminen, valvonta ja mahdolliset seuraamukset.

Tekijänoikeudet kouluissa

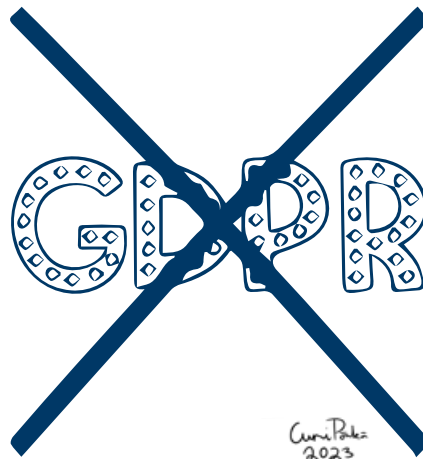
Opettajat voivat käyttää ainakin toistaiseksi tekoälysovellusten luomia kuvia ja tekstejä materiaaleissaan vapaasti, sillä tekoälyn tuotoksilla ei ole tekijänoikeus-

suoja. Lisäksi uuden tekijänoikeuslain opetuspoikkeus antaa oppilaille ja opettajille kouluympäristössä mahdollisuuden käyttää digitaalista materiaalia opetustarkeuksiin aiempaa laajemmin⁷¹.

Uusien sovellusten käyttöön liittyy kuitenkin merkittäviä eettisiä ja juridisia kysymyksiä, joista laajempia ratkotaan kansainvälisellä ja kansallisella tasolla. Opetushallituksen julkaiseman blogin mukaan⁷², lainsäädännön ja kunkin koulun ja oppilaitoksen omien sääntöjen mukainen, vastuullinen toiminta on tärkeää kaikessa digilaitteiden ja sovellusten käytössä, niin myös tekoälyyn tutustuttaessa. Tekoälyyn liittyvä osaaminen on osa tämän päivän ja tulevaisuuden yleissivistystä ja kuuluu kaikille oppijoille.

Lähdekritiikki

Tekoälyn antamista vastauksista on vaikea päätellä, mihin aineistoon ne perustuvat lähdetietojen puuttumisen vuoksi. On tapauksia, joissa on todettu tekoälyn keksineen lähdemateriaaleja, joita ei ole oikeasti olemassa. Oppilaiden ja opiskelijoiden kanssa töitä tehdessä ja tekoälysovelluksia käyttäessä kannattaa muistaa kriittinen suhtautuminen AI:n antamiin lähteisiin ja osa tehtävää tulisikin olla lähteiden tarkistaminen. Se on toki aikaavievää, mutta opettaa lähdekritiikkiä ja tiedonhakua perusteellisella tavalla.



Tietoturva

ChatGPT ja monet muut sovellukset eivät noudata ainakaan tätä kirjoittaessa EU:n tietosuojasetusta (GDPR). Siksi on tärkeää muistaa, ettei näille alustoille voi jakaa mitään luottamuksellista materiaalia esim. oppilastiedoista. GPT-järjestelmät saattavat olla myös ristiriidassa tietosuojasetuksen muiden kohtien kanssa, jonka pohjalta ihmisillä on esim. oikeus tulla unohdetuksi. Tietojen poistaminen GPT-mallista on kuitenkin ainakin tällä hetkellä mahdotonta.



CumPoka
2023

Läpinäkyvyyden puute

Markkinoille on tullut viimeisten kuukausien aikana lukuisia generatiivista tekoälyä hyödyntäviä ohjelmistoja ja sivustoja. Suurin osa niistä on kaupallisia ja kaikki niistä eivät ole välttämättä turvallisia ja luotettavia esim. tietosuojan kannalta. Monet organisaatiot ovat rajoittaneet uusien ohjelmistojen käyttöä odottaen tarkempia tietoturvaselvityksiä niiden toiminnasta. Koulujen kannalta tilanne on hankala. Harvalla koulutuksenjärjestäjällä on tarvittavaa tietotaitoa tai resursseja tarkistaa uusien ohjelmistojen mahdollisia riskejä.

Generatiivisen tekoälyn (GenAI) käyttämät keinotekoiset neuroverkot ovat yleensä "mustia laatikoita", eli niiden sisäistä toimintaa ei voida tutkia. Tämän vuoksi ne eivät ole "läpinäkyviä" eikä ole mahdollista selvittää, miten niiden tuotokset on syntyneet. Tekoälysovellukset voivat siis tuottaa odottamattomia tuloksia. Tämä läpinäkymättömyys on myös keskeinen syy GenAI-malleihin liittyviin luottamusongelmiin.⁷³ Jos käyttäjät eivät ymmärrä, miten GenAI-järjestelmä on päätenyt tiettyyn tulokseen, he eivät todennäköisesti ole halukkaita ottamaan sitä käyttöön tai käyttämään sitä.⁷⁴

Siksi kaikkien tulisi olla tietoisia siitä, että GenAI-järjestelmät toimivat kuin mustat laatikot ja että sen vuoksi on vaikeaa, ellei mahdotonta, tietää, miksi tietty sisältö on luotu. Tuotokset voivat myös heijastaa tiettyjä kulttuurisia tai kaupallisia arvoja, jotka vääristävät tuotettua sisältöä. Nykyisellään vaikuttaa siltä⁷⁵, että luotettavinta tietoa GPT-kielimallien opetusaineistoissa on Wikipedian sisältö, mutta sekään ei ole täysin neutraalia tai vapaata vinoumista⁷⁶, jotka periytyvät sen käyttäjäkunnalta.

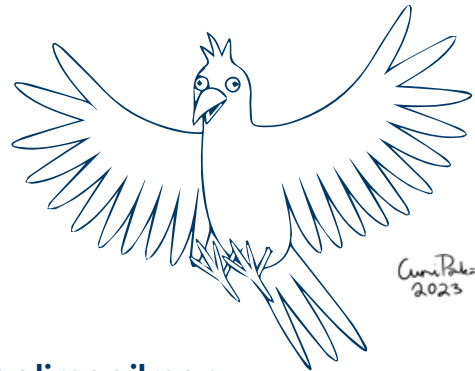
Virheellisen informaation leviäminen

GPT on koulutettu internetistä peräisin olevasta materiaalista, jossa on mukana usein disinformaatiota ja syrjivää tai muuta arveluttavaa sisältöä. GPT:n kehittäjät ovat pyrkineet niitä jossain määrin karsimaan, mutta tehokkaiden valvontamekanismien ja säännösten puuttuessa GenAI:n tuottama vinoutunut tai vääristynyt aineisto leviää eteenpäin. GenAI-järjestelmät saattavat siis tuottaa ja leviittää eteenpäin loukkaavaa ja epäeettistä materiaalia.

Vaikka tekoäly voi tuottaa tuloksia, jotka näyttävät huomattavan oivaltavilta ja hyödyllisiltä, se voi myös keksiä "tosiasioita", jotka kuulostavat täysin uskottavilta ja liittävät ne tuotoksiinsa. Tätä ilmiötä kutsutaan usein hallusinaatioksi⁷⁷. Virheet voivat olla syvälle kietoutu-

neita tekoälyn tuotoksiin, ja niitä voi olla vaikea havaita. Ne esiintyvät yleisimmin, kun kysytään lainauksia, lähteitä, viittauksia tai muita yksityiskohtaisia tietoja. Kannattaa siis aina tarkistaa annetut lähdetiedot ja viittaukset.

Vaikka GenAI:n tuottama materiaali voi vaikuttaa varsin tarkalta, ammattimaiselta ja vakuuttavalta, niin se voi sisältää virheellistä informaatiota ja vinoutuneita ajatuksia. Tämä on erityisesti suuri riski nuorille oppijoille, joilla ei ole vankkaa yleis- tai ennakkotietoa kyseisestä aiheesta.



CumPoka
2023

Reaalimaailman ymmärtämättömyys

Tekoälyn tuottama vastaus voi vaikuttaa usein niin luontevalta, että tuntuu kuin se ymmärtäisi tuottamansa tekstin sisältöä. GPT ei kuitenkaan ymmärrä tuottamaansa tekstiä ja se tuottaa usein virheellisiä lausumia. Siksi kaikkeen tekoälyn tuottamaan materiaaliin on suhtauduttava kriittisesti.

Tekoälyjärjestelmien tekstituotanto pohjautuu laskennallisiin todennäköisyyksiin, eikä niillä ole ymmärrystä generoiduista teksteistä. Siksi GenAI sovelluksia kutsutaan joskus "stokastisiksi papukaijoiksi." Vaikka ne voivat tuottaa vakuuttavalta tuntuvaa tekstiä, niin eivät ymmärrä, mitä sanovat. Ne toistavat harjoitusaineistostaan (yleensä Internetistä poimittua tekstiä) löytämiään kielellisiä malleja aivan kuten papukaija voi matkia ääniä ymmärtämättä todellisuudessa, mitä se sanoo.

Unescon raportin mukaan⁷⁸ GenAI:n perustana eivät ole havainnot todellisesta maailmasta tai tieteellisten menetelmien avulla hankitut tulokset, eikä se ole linjassa ihmisten tai sosiaalisten arvojen kanssa. Näistä syistä se ei voi tuottaa aidosti uutta sisältöä reaali maailmasta, esineistä ja niiden välisistä suhteista, ihmisistä ja sosiaalisista suhteista, ihmisen ja esineen välisistä suhteista tai ihmisen ja teknologian välisistä suhteista.

GenAI:ta tuottavat yritykset myöntävät, että GPT:t keksivät asioita, joita ei ole olemassa. Esimerkiksi ChatGPT:n julkisen käyttöliittymän alaosassa todetaan seuraavaa: "ChatGPT saattaa tuottaa epätarkkoja tietoja ihmisistä, paikoista tai tosiasioista".

Vähemmistöjen marginalisointi

GenAI-mallien kehittäjillä ja tarjoajilla on ensisijainen vastuu siitä, että näiden mallien tuotoksiin sisältyviä ennakkoluuloja käsitellään asianmukaisesti. Käyttäjien on kuitenkin hyvä olla tietoinen siitä, että GenAI-tekstin tuotos edustaa usein yleisintä tai vallitsevaa näkemystä maailmasta, sillä hetkellä, kun sen harjoitusaineisto tuotettiin, ja että osa siitä on ongelmallista tai sisältää ennakkoluuloja (esim. stereotyyppiset sukupuoliroolit)⁷⁹. Kielimalleista esimerkiksi ChatGPT on päässyt aiemmin käsiksi vain loppuvuoteen 2021 ulottuviin tietoihin, mutta syksyllä 2023 se ja vastaavat tekoälysovellukset ovat päässeet käsiksi reaaliaikaisiin tietoihin verkossa.

ChatGPT:llä ja vastaavilla työkaluilla on taipumus tuottaa vakiovastauksia, joissa heijastetaan mallien kouluttamiseen käytettyjen tietojen omistajien/luojien arvoja.⁸⁰ Jos sanasarja esiintyy usein harjoitusaineistossa - kuten yleisten ja kiistattomien aiheiden ja valtavirran tai vallitsevien uskomusten kohdalla - GPT toistaa sen todennäköisesti tulosteessaan. Tämä saattaa rajoittaa ja heikentää moniarvoisten mielipiteiden ja moniarvoisten ajatusten ilmaisujen kehittymistä. On olemassa väestöryhmiä, joilla on minimaalinen tai rajallinen digitaalinen läsnäolo verkossa. Heidän äänensä ei näin ollen tule kuulluksi eikä heidän huolenaiheitaan esitetä GPT:iden kouluttamiseen käytetyissä tiedoissa. GPT-mallit voivat siis entisestään marginalisoida jo ennestään heikommassa asemassa olevia ihmisiä. Siksi on hyvä tiedostaa, että vähemmistöjen äänet voivat jäädä vastauksissa huomiotta.

Ikärajasuositukset

Generatiivisen tekoälyn koulukäytössä on huomioitava ohjelmistojen antamat ikäsuositukset. Esim. OpenAI ilmoittaa sivuillaan⁸¹, ettei ChatGPT:tä ole tarkoitettu alle 13-vuotiaille lapsille, ja että 13-18-vuotiaiden lasten tulisi hankkia vanhempiensa suostumuksen ennen ChatGPT:n käyttöä, sillä ohjelmisto saattaa tuottaa tulkoksia, joka ei sovellu kaikille yleisöille tai kaikenikäisille. Opettajien tulisi ottaa tämä huomioon käyttäessään ohjelmistoa oppilaiden kanssa. Lisäksi jos ChatGPT:tä käytetään alle 13-vuotiaiden oppilaiden opetuksessa, niin vuorovaikutuksen ChatGPT:n kanssa on oltava aikuisen suorittama.

Ympäristövaikutukset

Uusien digitaalitekniikoiden käyttöön liittyy läheisesti myös ympäristön ekologista kestävyyttä liittyviä näkökohtia. Generatiivisen tekoälyn ympäristökustannukset jätetään usein huomiotta. Tekoälyn kouluttaminen ja Bitcoinin kaltaisten kryptovaluuttojen louhiminen ovat erittäin resurssi-intensiivisiä prosesseja datan ja lasken-

tatehon osalta.

Datakeskukset vastaavat noin 2-3 prosentista maailman kasvihuonekaasupäästöistä⁸². Datan määrä maailmassa kaksinkertaistuu joka toinen vuosi. Tietokeskusten palvelimet, jotka tallentavat tätä jatkuvasti kasvavaa tietomerta, vaativat valtavia määriä energiaa ja vettä (suoraan jäähdytykseen ja epäsuorasti uusiutumattoman sähkön tuottamiseen) tietokonepalvelimien, laitteiden ja jäähdytysjärjestelmien käyttämiseen. Suurella energiankulutuksella on huomattavia ympäristövaikutuksia.

DigComp 2.2:ssa suositellaan tekoälyn käyttäjiä⁸³ pohtimaan tekoälyjärjestelmien käytön eettisiä seurauksia koko niiden elinkaaren ajan. Näihin kuuluvat sekä ympäristövaikutukset (digitaalisten laitteiden ja palvelujen tuotannon ympäristövaikutukset) että yhteiskunnalliset vaikutukset, jotka voi rajoittaa työntekijöiden yksityisyyttä tai oikeuksia esimerkkinä halpatyövoiman käyttö kuvien merkitsemiseen tekoälyjärjestelmien kouluttamiseksi.

Toisaalta tekoälyä voidaan käyttää myös ympäristöystävällisiin tarkoituksiin esim. kehittämällä järjestelmiä, jotka optimoivat erilaisten tilojen energiatehokkuutta seuraamalla lämmitystarvetta ja ilmastointia.

Kuluttajien on hyvä olla tietoinen teknologian ympäristövaikutuksista, jotta he osaisivat valita ekologisia ja kestävästi kehitystä edistäviä digitekniikan vaihtoehtoja ja toimintatapoja.

Laadukkaat väärennökset

Generatiivista tekoälyä hyödyntäviä ohjelmistoja voidaan käyttää kuvien tai videoiden muuttamiseen tai manipulointiin. Näitä deepfake-väärennöksiä on vaikea erottaa aidoista kuvista tai videoista. Aikaisemmin huijaussivustot ja disinformaatiota sisältävät sivustot saatiin tunnistaa amatöörimäisesti tehdyistä videoista ja käännösvirheitä sisältävistä kömpelöistä teksteistä.

Nykyään disinformaation levittäjien on GenAI:n avulla helppo luoda erittäin aidon näköisiä väärennöksiä, joissa voidaan käyttää luvatta esim. jonkun kasvoja ja ääntä. GenAI helpottaa näin myös disinformaation levittämistä, vihapuheen lisääntymistä ja valitettavasti myös rikollista toimintaa. Tekoäly voi auttaa tietoverkko-rikollisia tuottamaan haittaohjelmia nopeasti, automaattisesti hyökkäyksiä ja tehostamaan huijauksia tai sosiaalista manipulointia koskevia hyökkäyksiä käyttämällä väärennöksiä ja tekoälyllä toimivaa ihmisen kulloista äänisynteesiä⁸⁴.

Esimerkkitapaus, BBC 16.9.2023⁸⁵

Lapsille suunnatut suositut ”opetussisällöt” sisältävät tekoälyn avulla vääristeltyä ”tieteellistä” informaatiota, videoita ja sisältöjä.

BBC:n Global Disinformation Team löysi YouTubesta kymmeniä kanavia, jotka tuottavat harhaanjohtavaa pseudo-tieteellistä materiaalia lapsille muun muassa arabiaksi, espanjaksi ja thaiksi. Monilla näistä kanavista oli yli miljoona tilaajaa. Niiden videot saavat usein miljoonia katselukertoja.

Kanavien tekijät julkaisevat sisältöä nopeasti, ja monet julkaisevat useita videoita päivittäin. Pystyäkseen tekemään näin nopeasti BBC:n toimittajat epäilivät, että sisältötuottajat käyttivät generatiivisia tekoälyohjelmia. Nämä ovat Chat GPT:n ja MidJourneyn kaltaisia ohjelmia, jotka voivat luoda uutta sisältöä, kun niitä pyydetään [esim. ”musta kissa, jolla on kruunu päässään”], sen sijaan että etsisivät internetistä jo olemassa olevia esimerkkejä.

Tämän teorian testaamiseksi he ottivat videoita jokaiselta kanavalta ja käyttivät tekoälyn tunnistustyökaluja ja asiantuntija-analyysiä arvioidakseen todennäköisyyttä tai todennäköisyyttä sille, että kuvamateriaali, selostus ja käsikirjoitus oli tehty tekoälyn avulla. BBC:n analyysi osoitti, että useimmissa videoissa tekoälyä oli käytetty tekstin ja kuvien tuottamiseen sekä tietojen poimimiseen verkkosivustolta ja materiaalin manipulointiin todellisista tiedevideoista. Tuloksena on sisältö, joka näyttää tosiseikoilta, mutta on useimmiten valheellista.

⁸⁷ European Commission (2023) International Draft Guiding Principles for Organizations Developing Advanced AI systems.

<https://digital-strategy.ec.europa.eu/en/library/international-draft-guiding-principles-organizations-developing-advanced-ai-systems>

⁸⁸ Council of EU (2023). Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world.

<https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>

⁸⁹ Cloudware (ladattu 19.12.2023) What is robots.txt? | How a robots.txt file works. <https://www.cloudflare.com/en-gb/learning/bots/what-is-robots-txt/>

⁹⁰ Kopioisto (ladattu 19.12.2023) Kopiraitilla. <https://kopiraitilla.fi/tekijanoikeustietoa/opettajan-tekijanoikeus/>

⁹¹ Opetushallitus (ladattu 18.12.2023) Tekoälyä hyödyntäviin sovelluksiin tutustuminen koulussa ja oppilaitoksissa.

<https://www.oph.fi/fi/opettajat-ja-kasvattajat/joko-teidan-tyoyhteisossanne-keskustellaan-tekoalysta-0>

⁹² Nazaretsky, T., Cukurova, M. and Alexandron, G. 2022a. An Instrument for Measuring Teachers' Trust in AI-Based Educational Technology. LAK22: 12th International Learning Analytics and Knowledge Conference. Vancouver, Association for Computing Machinery, pp. 55-66.

⁹³ Nazaretsky, T. et al (2022) Teachers' trust in AI-powered education technology. British Journal of Educational Technology. BERA. <https://doi.org/10.1111/bjet.13232>

⁹⁴ Laaksonen, S. (2023) Tekstiä tuottava tekoäly ja tieteen avoimuus. <https://blogs.helsinki.fi/thinkopen/tekoaly-ja-tieteen-avoimuus/>

⁹⁵ Wikipedia (ladattu 19.12.2023) https://en.wikipedia.org/wiki/Wikipedia:Systemic_bias

⁹⁶ Mollick, E.& Mollick, L. (2023) Assigning AI: seven approaches for students, with prompts. <http://dx.doi.org/10.2139/ssrn.4475995>

⁹⁷ Fengchun, M. (2023). Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

⁹⁸ Fengchun, M. (2023). Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

⁹⁹ Ks. esim. Coded bias-dokumenttifiili: <https://www.netflix.com/es-en/title/81328723>

¹⁰⁰ OpenAI (ladattu 19.12.2023) Is ChatGPT safe for all ages? <https://help.openai.com/en/articles/8313401-is-chatgpt-safe-for-all-ages>

¹⁰¹ Kumar, A., and Davenport, T. (2023) How to make generative AI greener. HBR. <https://hbr.org/2023/07/how-to-make-generative-ai-greener>

¹⁰² Vuorikari, R., Kluzer, S. and Punie, Y., (2022) DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg. doi:10.2760/490274

¹⁰³ Belani, G. (2023) AI for Cybersecurity and Cybercrime. IEEE Computer Society. <https://www.computer.org/publications/tech-news/trends/ai-fighting-ai>

¹⁰⁴ BBC (2023) AI used to target kids with disinformation. <https://www.bbc.com/newsround/66796495>



8. Tekoöly haastaa luku- ja kirjoitustaidon

HARTO PÖNKÄ

Voisi luulla, että tekoöly on luku- ja kirjoitustaidottoman suurin ystävä. Niin tavallaan onkin, sillä hakeehan generatiivinen tekoöly nopeasti vastauksen lähes mihin tahansa kysymykseen sekä luo pyynnöstä vaikka kokonaisen artikkelin.

Samaan aikaan juuri tässä piilee yksi tekoölyn suurimmista haasteista: jos totumme kysymään asioita tekoölysovelluksilta sen sijaan, että hakisimme tietoa

itse ja arvioisimme löytämäämme tietoa kriittisesti, saatamme näivettää lukutaitomme. Sama koskee erilaisten tekstien tuottamista ja kirjoitustaitoa.

Sanomattakin on selvää, että jos tekoölysovelluksia tottuu käyttämään tiedonhaussa ja tekstien tuottamisessa jo peruskouluvaiheessa, voivat kehittyneet luku- ja kirjoitustaidot jäädä alun perinkin oppimatta.

Tekoäly voi heikentää lasten ja nuorten lukutaitoa entisestään

Kun jokin uusi teknologia kehittyy siihen pisteeseen, että sen käyttö yleistyy nopeasti, ottavat lapset ja nuoret sen muiden joukossa käyttöön. Siksi voidaan odottaa, että he omaksuvat pian myös tekoälysovellusten käytön. Jo tällä hetkellä monissa nuorten suosimissa sovelluksissa on tekoälyominaisuuksia. Esimerkiksi Snapchat-pikaviestisovelluksessa oli ensimmäisten joukossa OpenAI-yrityksen ChatGPT:hen perustuva tekoälykaveri.

Tekoälysovellusten suosio perustuu osaltaan helpokäyttöisyyteen. Niiden käytön aloittaminen on niin yksinkertaista, että lapsikin sen taitaa. Tämän takia voisi luulla, että lapset ja nuoret ovat taitavia tekoälysovellusten käyttäjiä, ja että he pystyisivät huomioimaan sekä välttämään tekoälysovelluksiin liittyviä ongelmia. Tällainen luulo ei kuitenkaan kestä lähempää arviointia.

Valitettavasti minkään teknologian runsas käyttö ei takaa, että käyttäjä huomaisi taitavasti siihen liittyviä ongelmia. Pikemminkin hänen voi odottaa tottuvan mahdollisiin ongelmiin, ikään kuin hyväksyvän ne ja tulevan niille sokeaksi.

Uudet ilmiöt kiinnostavat nuoria luonnostaan, ja houkutus helpottavien työkalujen käyttöön lienee selvä. Lapset ja nuoret tarvitsevat kuitenkin apua siihen, miten ja milloin tekoälysovelluksia kannattaa käyttää, sekä miten arvioida niiden antamia vastauksia. On täysin eri asia kysyä generatiiviselta tekoälyltä vastauksia kysymyksiin kuin hakea itse tietoa eri lähteistä ja arvioida sitä kriittisesti. Tekoälysovellukset ovat omiaan vähentämään lukemista ja eri lähteiden vertailua.

Vasta-argumenttina joku voisi esittää, että tekoälysovellusten ansiosta niiden käyttäjät lukevat aiempaa enemmän, sillä antavathan sovellukset vastauksena tekstiä. Tiedonhaun näkökulmasta näin ei kuitenkaan ole, sillä tavallisesti haemme tietoa eri lähteistä – tai ainakin meidän pitäisi tehdä niin – ja rakennamme käsityksemme luotettavina pitämiemme tekstien pohjalta sen sijaan, että saisimme ns. valmiiksi pureskellun synteessin sovellukselta. Generatiivisten tekoälysovellusten ominaisuuksiin suorastaan kuuluu valtavan lähde-

tietomassan tiivistäminen käyttäjälle mukavan lyhyeksi tulokseksi.

On perusteltua odottaa, että tekoälysovellusten yleistyminen vähentää lasten ja nuorten lukemista entisestään. Luetaanko edes tekoälyn antamia vastauksia pian enää kriittisesti?

ChatGPT ja Kalle Päätalon vaimot

Otetaan esimerkki. Kirjailija Kalle Päätalo on henkilö, jonka ChatGPT-tekoälyn voi odottaa tuntevan hyvin – onhan hänestä runsaasti tietoa lukemattomilla nettisivustoilla. Samalla hän on kirjailija, jota voidaan odottaa käsiteltävän esimerkiksi lukion äidinkielen opetuksessa.

Syksyllä 2023 törmäsin Kalle Päätalon vaimoja koskevaan kysymykseen uudessa Bluesky-somepalvelussa. Eräs keskustelija toi esiin, että hän oli kysynyt ChatGPT:itä, keiden kanssa Päätalo oli ollut naimisissa. Kysyin itse ChatGPT:itä kokeeksi samaa ja sain oheisen vastauksen.

Vastaus on ensilukemalta täysin uskottava. Kalle Päätalon vaimojen nimet ja avioliittojen vuosiluvut ovat yksityiskohtia, joissa olevia virheitä ei voi huomata muuten kuin tarkistamalla ne jostain muusta lähteestä. Siksi ChatGPT:n käyttäjät ovat hyvin alttiita tämän kaltaisille virheille.

Tarkistin, mitä Päätaloo käsittelevä Wikipedia-artikkeli¹⁰⁵ aiheesta sanoo. Sen mukaan ChatGPT:n vastaus on täysin väärin. Wikipedian mukaan Päätalo avioitui Helvi Ojalan (Päätalon teoksissa nimeltään Laina Puronen) kanssa 28.5.1944 ja Elli Helena (Leena) Janakan kanssa 14.9.1955. Wikipedia-artikkelissa ei ollut yksityiskohtaisia lähteitä näille tiedoille, mutta nimien oikeellisuus oli helppo tarkistaa hakemalla lisätietoa muista lähteistä.

Pyysin ChatGPT:tä tarkistamaan vastauksensa neljästi. Joka kerralla sekä vaimojen nimet että vuosiluvut olivat väärin. Vain avioliittojen lukumäärä ja Päätalon kuolinvuosi osuivat oikeaan. Lopulta ChatGPT luovutti ja pyysi tarkistamaan tiedot luotettavista lähteistä. Tätä neuvoa voi pitää hyvänä.



You

Keiden kanssa Kalle Päätalo oli naimisissa?



ChatGPT

Kalle Päätalo oli suomalainen kirjailija, ja hän oli naimisissa kahdesti. Ensimmäinen vaimo oli Aili Eerolan kanssa, ja heidän avioliittonsa kesti vuosina 1949–1966. Toisen kerran Päätalo meni naimisiin Anna-Liisa Haaviston kanssa vuonna 1966, ja heidän liittonsa jatkui Päätalon kuolemaan asti vuonna 2000.



ChatGPT

Pahoittelen suurta sekaannusta ja aiheuttamaani pettymystä vastauksissani. Jos minulla on edelleen virheellisiä tietoja, suosittelen tarkistamaan Kalle Pääatalon perhesuhteita luotettavista lähteistä, kuten virallisista biografoista tai haastatteluista. Olen pahoillani mahdollisista väärinkäsityksistä ja kiitän kärsivällisyydestäsi.



Bluesky-somepalvelussa keskustelu Kalle Pääatalosta jatkui edelleen niin, että eräs toinen keskustelija kertoi saaneensa ChatGPT:ltä oikean vastauksen käyttämällä palvelun maksullista Plus-palvelua. Plus-palvelussa ChatGPT:tä voi pyytää tarkistamaan tiedot internetistä.

Ajatus siitä, että tekoälyä tulee pyytää erikseen tarkistamaan antamansa tiedot, tuo hyvin esiin generatiiviseen tekoölyyn liittyviä ongelmia. Voidaanko tekoälyä ylipäätään käyttää jonkin asian tarkistamiseen?

Toiseksi voidaan kysyä, oliko ChatGPT alun perinkin järkevä työkalu Kalle Pääatalon vaimojen tietojen hakuun. Tarkalleen ottaen ChatGPT on laajaan kieli-malliin pohjautuva keskustelubotti, joka tuottaa keho- teeseen todennäköisesti liittyvän tekstivastauksen. Se ei siis ole hakukone, joten tässä tapauksessa se oli väärä työkalu tiedonhakuun. Tämä selittää sen, miksi ChatGPT tekee virheitä, mutta osaa samalla olla ystävällinen ja tarjota yleispäteviä neuvoja.

Varmimmin oikean tiedon olisi löytänyt käyttämällä netin hakukoneita tai aiheeseen sopivaa lähde- teosta. Syy, miksi joku käyttäisi tällaiseen tiedonhakuun Chat- GPT:tä voisi olla tottumus saada siltä nopeasti vastauksia sen sijaan, että näkisi itse tiedonhaun vaivan. Tässä tapauksessa ChatGPT:n käyttö johti kuitenkin väriin tietoihin ja lisävaivaan tietojen tarkistamisessa.

Tekoälytyökalut voivat olla hyödyllisiä myös tiedon- haussa, mutta silloin tulee käyttää nimenomaan haku- koneeksi tarkoitettua sovellusta. Sellainen on esimer- kiksi Microsoftin Bing-hakukoneeseen liitetty Bing Chat. Se olisi tässä tapauksessa osannut kertoa toisen Kalle Pääatalon vaimon nimen oikein sekä tarjonnut verkko- lähteitä tiedon tarkistamiseen.

Tekeekö generatiivinen tekoäly meistä kaikista hujareita?

Opetuksen ja teknologian trendejä luotaava tuorein Educause Horizon¹⁰⁶ -raportti nostaa suomalaisesta- kin tekoälykeskustelusta tutut teemat esiin. Kehittyneet generatiiviset ja ennustavat tekoälymallit ovat teknolo- gia, jolla tulee olemaan lyhyellä ja pitkällä aikavälillä merkittäviä muutoksia niin opiskeluun, työelämään kuin yhteiskunnalliseen osallistumiseenkin.

ChatGPT:n ja muiden generatiivisten tekoälysovellus- ten odotetaan Educausen raportin mukaan johtavan siihen, että opiskelijat laittavat koneen tekemään tehtä- vät puolestaan. Oppija voi esimerkiksi pyytää tekoälyä kirjoittamaan esseen tai ratkaisemaan oppikirjasta va- lokuvatut tehtävät puolestaan. Tähän pystyvät sovel- lukset ovat jo nyt joko ilmaisia tai hyvin edullisia.

Tehtävissä huijaamisen paljastamiseen kehitetään luonnollisesti jo nyt uusia ohjelmia opettajien avuk- si. Kyse on eri tekoälysovellusten kehittäjien välisestä kilpajuoksusta, joka ei kuitenkaan ratkaise varsinaista ongelmaa.

Tällä hetkellä generatiiviset tekoälysovellukset te- kevät vielä kohtalaisen helposti tunnistettavaa tekstiä, mutta hyvin pian tilanne muuttuu. Sen jälkeen tekoä- lyn tuottaman tekstin erottaminen ihmisen tuottamasta tekstistä on käytännössä mahdotonta. On odotettavis- sa, että kielimallit oppivat tekemään myös inhimillisel- tä näyttäviä virheitä tai simuloimaan tietyn tyyppistä kirjoittajaa niin pyydettyä.

Houkutus tekoälysovellusten käyttöön tekstien kir- joittamisessa ei koske vain koululaisia ja opiskelijoita. Generatiivista tekoälyä käytetään jo esimerkiksi some-

Keiden kanssa Kalle Pääatalo oli naimisissa?

✓ Haetaan kohdetta: **Kalle Pääatalo avioliitto**

✓ Luodaan vastauksia sinulle...

Kalle Pääatalo oli naimisissa kahdesti. Ensimmäisen avioliiton jälkeen hän avioitui Leena Janakan kanssa vuonna 1955. Helle syntyi kakai tytärtä, Riitta ja Elina ^{1 2 3}.

Lisätietoja: [1. bing.com](https://1.bing.com) [2. taivalkonki.fi](https://2.taivalkonki.fi) [3. kirjojentakana.fi](https://3.kirjojentakana.fi) [4. is.fi](https://4.is.fi) [5. tampere.fi](https://5.tampere.fi)

1 / 5

julkaisujen, työhakemusten, myyntikirjeiden, konferenssi-esitysten, uutistekstien, disinformaation, runojen ja laulusanoitusten generoimiseen.

Vaikka tekoälysovellusten tuottama teksti näyttäisi uskottavalta, siinä piilee vaaransa. Ensinnäkin se, että osaamme pyytää tekoälyä kirjoittamaan tietystä aiheesta, ei opeta meille itse aiheesta välttämättä mitään. Neuroverkkoihin tallentuneet eri sanojen väliset yhteydet eivät siirry ilman oppimista omaan muistiimme – saati, että ymmärtäisimme, miten ja miksi asiat liittyvät toisiinsa.

Educausen raportti tuo esiin korkeakoulutasolla asti opettajilla olevan huolen, että opiskelijat luottavat liikaa tekoälyn antamiin vastauksiin. Tekoälyn käytössä on huomioitava niiden rajoitukset ja eettiset vaikutukset, jotka liittyvät esimerkiksi tekoälyn antamien vastausten vinoumiin. Ilman näiden huomiointia tekoälytyökalujen käyttö voi altistaa värien tietojen omaksumiseen ja tekoälyn epäasialliseen käyttöön.

Taito osata kirjoittaa hyviä prompteja eli kehoitteita tekoälylle ei korvaa jonkin asiakokonaisuuden syvällistä ymmärtämistä tai johonkin ammattiin liittyvää niin sanottua hiljaista tietoa. Tämä tulee tekoälysovellusten käyttäjille vastaan viimeistään silloin, kun tekoälyltä kysytään asiaa, johon se ei yksinkertaisesti osaa vastata. Generatiivinen tekoäly osaa kertoa asiasta kuin asiasta yleistiedot, mutta jos siltä kysyy esimerkiksi jonkin ammattialan sisäiseen sanastoon kuuluvan käsitteen merkitystä, jää vastaus todennäköisesti saamatta tai se on

auttamattoman puutteellinen.

Esimerkiksi lyhenne KÄTS (joka tulee sanoista kirjain, äänne, tavu, sana) lienee tuttu jokaiselle peruskoulun alaluokkien opettajille, mutta ChatGPT:ltä siitä on turha kysyä. Sovellus sentään osaa antaa neuvon, mistä haakea aiheesta tietoa.

Tottuminen tekstin tuottamiseen tekoälymallien avulla voi korvata toisen paljon tärkeemmän taidon: oman kirjoitustaidon. Ilman sujuvaa kykyä ilmasta ajatuksiimme tekstimuodossa, emme pysty osallistumaan moniin arkipäiväisiin viestintä- ja yhteistyötilanteisiin, jotka tapahtuvat teknologian välityksellä. Kirjoitustaidon aktiivinen käyttäminen on yhteydessä muihinkin ajattelun taitoihin. Jos emme pysty ilmaisemaan kielellisesti, miten ymmärrämme asioita, emme itse asiassa ymmärrä niitä kovin syvällisesti.

Lukutaidon merkitys korostuu tekoälyn aikakautena

Lyhyellä aikavälillä on odotettavissa, että kokonaan tekoälyllä tai sen avulla tuotettujen tekstien määrä kasvaa varsinkin netti- ja sosiaalisen median palveluissa räjähdysmäisesti. Tekstien lisäksi koneoppimismallien avulla tuotetaan kuvia ja videoita. Tekoälyn avulla tuotettuja sisältöjä käytetään niin tavanomaisissa ja kaupallisissa somesisällöissä kuin tarkoituksella tehdyssä disinformaatiossa.



You

Mitä tarkoittaa lyhenne KÄTS lukemisen opetuksessa?



ChatGPT

Valitettavasti minulla ei ole tietoa lyhenteestä "KÄTS" lukemisen opetuksen kontekstissa, ainakaan tietoni leikkauspisteeseen tammikuu 2022 asti. Lyhenteet voivat vaihdella eri opetusohjelmien ja maiden välillä, ja niiden merkitys voi olla paikallisesti määritelty.

Jos "KÄTS" liittyy johonkin tiettyyn opetussuunnitelmaan, opetusmenetelmään tai koulutusohjelmaan, suosittelen tarkistamaan paikallisen opetussuunnitelman, koulun materiaalit tai opetusviranomaisten verkkosivustot. Saatat löytää sieltä tarkempaa tietoa siitä, mitä lyhenne tarkoittaa ja miten sitä käytetään kyseisessä kontekstissa.



Kuvio 2.7 Lukutaidon keskiarvot vuosina 2000–2022 Suomessa ja eräissä vertailumaissa



Jo ennen nykyistä tekoälybuuria suomalaisnuorten lukutaidon on todettu PISA-tutkimuksissa heikentyneen¹⁰⁷. Vuoden 2022 PISA-tulosten mukaan suomalaisnuorten lukutaito oli edelleen laskenut jyrkästi.⁸⁹ Lukutaidon heikentyminen näyttää olevan yhteydessä kirjojen lukemisen vähenemiseen, joka on yhteydessä siihen, että vapaa-ajasta yhä suuremman osan vie sosiaalisen median käyttö ja muu ruutu-aika. Somessa kyllä luetaan, mutta tekstit ovat aiempaa lyhyempiä.¹⁰⁸

Somepalvelujen käytön lisääntyminen ja lukutaidon heikentyminen muodostavat viheliäisen ja vaikeasti ratkaistavan ongelman: jos käytämme yhä enemmän aikamme sosiaalisen median sisältöjen parissa, luku- ja tiedonhakutaitomme heikkenee, vaikka juuri somessa tarvitsimme timanttisen terävää lukutaitoa.

Diginatiivi-käsitteellä viitataan nuoriin, jotka ovat viettäneet koko elinaikansa nykyisen tieto- ja viestintätekniikan ympäröimänä. Valitettavasti runsas somepalvelujen käyttö ei näytä takaavan diginatiiveille edes sosiaalisen median keskustelujen ja niissä esitettyjen väitteiden luetun ymmärtämistä. Tämä käy ilmi Kansallisen koulutuksen arviointikeskuksen eli KARVIN tutkimuksesta, jossa selvitettiin peruskoulun yhdeksäsluokkalaisten kirjoitus- ja lukutaitoja.¹⁰⁹

KARVIN raportin mukaan keskustelumuotoisen netitekstin argumentin ja perustelujen ymmärtäminen oli haasteellista jopa neljäsosalle 9.-luokkalaista, vaikka kyse oli heille tutusta tekstimaailmasta. Suurella osalla nuorista oli siten merkittäviä puutteita ymmärtää heille tutumpien lyhyiden sometekstien sisältämiä väitteitä. On sanomattakin selvää, että jos väitteitä ja niiden perusteluja ei ymmärretä, ei niitä myöskään pystytä arvioimaan kriittisesti. Raportti huomauttaa, että me-

diatekstien lukeminen on yhteiskunnallisen keskustelun ja kaupallisen vaikuttamisen näkökulmasta tärkeä kansalaistaito.

Useat nuorten käyttämät sosiaalisen median palvelut sisältävät runsaasti kaupallisessa tarkoituksessa tehtyä sisältöä. Näin on esimerkiksi Instagramissa ja YouTubessa. Disinformaation eli tarkoituksella levitetyn virheellisen informaation määrä on merkittävä niin X:ssä (ent. Twitter), Facebookissa, Instagramissa kuin TikToksissakin.¹¹⁰

Kriittistä lukutaitoa tutkivan Critical-ryhmän tutkimusten mukaan kuudesluokkalaista 40 % ei kyseenalaistanut lainkaan kaupallisen tekstin kirjoittajan asiantuntijuutta. Lisäksi suurin osa tutkimuksen osallistuneista nuorista ei osannut tulkita harhaanjohtavia diagrammeja.¹¹¹

Kun tekoälyllä tuotettujen tekstien, kuvien ja videoiden määrä kasvaa, kriittisen lukutaidon merkitys korostuu entisestään. Samalla tarvitaan uudentyyppisiä lukemisen taitoja, jotka liittyvät tekoälyllä tuotettujen tekstien ja muiden sisältöjen tunnistamiseen ja tulkitaan. Ilman kehittyneitä digitaalista informaatiolukutaitoa emme esimerkiksi osaa epäillä ja tutkia, milloin kyse on valeutisesta, tekoälymallin luomasta valokuvaa muistuttavasta kuvasta tai deepfake-videosta. Tekoälyyn liittyviä osaamistarpeita tässä julkaisussa käsittelee lähemmin Kari Kivinen artikkelissaan "Tekoäly - mitä meidän pitäisi tietää ja osata?".

Informaatiolukutaito voi rakentua vain hyvän peruslukutaidon ja kriittisen lukutaidon varaan. Tämän takia peruslukutaidon opetus on ensiarvoisen tärkeää myös silloin, kun opetamme tekoälyn vastuulliseen käyttöön liittyviä taitoja.

Miten sovittaa yhteen tekoälyn sekä luku- ja kirjoitustaitojen opetus?

Opetusalalla on havahduttu nopeasti pohtimaan, miten tekoälyä voidaan hyödyntää opetuksessa. Tässäkin teoksessa kerrotaan kokemuksia ja pedagogisia vinkkejä, miten varsinkin generatiivista tekoälyä voidaan käyttää opetuksessa. Esimerkiksi ChatGPT:tä voi käyttää koulutehtävissä ideoinnin apuna ja vieraiden kielten opiskelussa keskusteluharjoituksissa.

Luku- ja kirjoitustaidon opetuksen kehittämiseen keinoja tarjoaa jo edellä lainattu KARVIN tutkimus.¹¹² Se nostaa esiin muun muassa seuraavia suosituksia:

- Luetun ymmärtämisen harjoittelua tulee jatkaa vielä yläkoulussa luetun ymmärtämisen strategioiden ja lukemisen mallintamisen avulla.
- Sosiaalisessa mediassa tarvitaan väitteiden ja perusteluiden analysointia ja tulkintaa. Tällainen lukeminen vaatii tekstien kriittisen ja kontekstuaalisen tulkinnan taitoa eli sen tulkitsemista, mitä tekstien sisällöt merkitsevät juuri kyseisessä ympäristössä.
- Kirjoitustaitoja tulee lähestyä monilukutaito-ajattelun kautta. Rakentakaa monimuotoisia tekstejä netissä, monenlaisilla alustoilla ja työvälillä, pareittain, yksin ja ryhmissä.
- Jokainen opettaja on samalla kielenopettaja sekä ajattelun ja oppimaan oppimisen opettaja. Esimerkiksi opiskelijan oma tavoitteenasettelu, itsearviointi ja ongelmanratkaisu kuuluvat jokaiseen oppimiskokonaisuuteen.

Perusopetuksen opetussuunnitelman perusteissa (POPS) tekoälyä ei vielä mainita.¹¹³ Opetussuunnitelman perusteet eivät silti ole tältä osin vanhentuneet, sillä tekoälyllä tuotetut sisällöt liittyvät monilukutaidon laaja-alaiseen osaamiseen. POPSin mukaan oppijoita ohjataan tulkitsemaan, tuottamaan ja arvioimaan yhä moninaisempia tekstejä erilaisissa yhteyksissä ja ympäristöissä. Samalla harjoitellaan fiktion, faktan ja mielipiteen erottelua sekä huomaamaan tekstien erilaisia tavoitteita ja keinoja.

Ehkä tärkein tekoälyyn liittyvä monilukutaidon osa-alue on kyky vertailla ja arvioida haetun tiedon soveltuvuutta. Opetussuunnitelman perusteiden mukaan kriittistä lukutaitoa tulee kehittää oppijoille merkityksellisissä asiayhteyksissä. Opetuksessa kannattaa käsitellä ainakin yleisimpiä tekoälyvälineitä ja sellaisia tekoälyn käyttötarkoituksia, jotka ovat oppijoille tuttuja. Näitä ovat esimerkiksi juuri edellä käsitellyt sosiaalisen median sisällöt, ChatGPT ja tekoälyn tuottamien vastausten arvioiminen.

Monilukutaitoa digiaikaan -teoksen artikkelissa olen aiemmin kirjoittanut monilukutaidon merkityksestä sosiaalisessa mediassa.¹¹⁴ Monilukutaitoa arvioidessa huomio kiinnittyy oppijoiden kykyyn ymmärtää sisällön eri tasoja ja niihin liittyviä merkityksiä. Tämä korostuu entisestään tekoälyn avulla tuotettujen sisältöjen tulkinna. Taitavat tekoälyn toimintaa ymmärtävät oppijat kykenevät löytämään ja arvioimaan esimerkiksi seuraavia asioita:

- Onko sisältö ihmisen vai tekoälyn tekemä?
- Mitkä asiat sisällössä viittaavat tekoälyn hyödyntämiseen?
- Mitä tekoälysovellusta sisällön tekoon on saatettu käyttää?
- Onko olemassa työkaluja, joilla voitaisiin varmentaa, onko sisältö ihmisen vai tekoälyn tuottama?
- Mihin tekoälyn tuottama sisältö perustuu?
- Milloin sisältö on julkaistu? Onko kyse ensijulkaisusta vai onko sisältö julkaistu jo aiemmin muualla?
- Onko sisältö tiedollisesti oikein vai onko siinä virheitä tai vinoumia? Mitä niistä voidaan päätellä?
- Onko valittua tekoälysovellusta käytetty tarkoituksenmukaisesti?
- Onko sisältö tehty eettisesti vastuullisesti?
- Mitä tekijänoikeuksia sisältöön liittyy? Onko sen levittämisessä rikottu toisen tahon tekijänoikeuksia?
- Onko tekoälylle annettua kehoitetta kehitetty parhaan mahdollisen vastauksen saamiseksi?
- Voidaanko samaa asiaa kysyä joltain muulta tekoälysovellukselta? Tuottaisiko se todennäköisesti saman vai eri vastauksen?
- Mihin laajempaan sisältökokonaisuuteen tai ajankohtaiseen keskusteluun sisältö mahdollisesti liittyy?
- Mitä käytetty tekoälysovellus ja sisältö kertovat sen tekijästä?
- Mitä tarkoitusta varten sisältö on tehty ja mistä tämä on pääteltävissä?
- Ketkä ovat jakaneet sisältöä sosiaalisessa mediassa ja miten? Mitä tästä voidaan päätellä?

Oppimista ja syvällistä ymmärtämistä arvioitaessa kyse on siitä, kuinka monipuolisesti oppija tunnistaa ja osaa selittää sisältöön, sen tekijään ja sosiaaliseen kontekstiin liittyviä seikkoja, merkityksiä ja niiden välisiä yhteyksiä. Viime kädessä monilukutaidon hallinta näkyy kykyinä toimia taitavasti ja vastuullisesti erilaisissa ympäristöissä – myös tekoälysovellusten aikana.

105 Wikipedia, (2023). Kalle Päätalo. https://fi.wikipedia.org/wiki/Kalle_P%C3%A4%C3%A4t%C3%A4l%C3%B6

106 Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., Grajek, S., Birdwell, T., Liu, D., Mandernach, J., Moore, A., Porcaro, A., Rutledge, R., ja Zimmern, J. (2023). EDUCAUSE Horizon Report, Teaching and Learning Edition. <https://library.educause.edu/-/media/files/library/2023/4/2023hrteachinglearning.pdf>

107 Opetus- ja kulttuuriministeriö, (2019). Pisa 18 ensituloksia. <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161922/Pisa18-ensituloksia.pdf>

108 Hanifi, Riitta (17.3.2022). Lapset lukevat aiempaa vähemmän kirjoja - netissä ja somessa lukutaito määrittyy uudelleen. Tilastokeskus. <https://www.stat.fi/tietotrendit/artikkelit/2022/lapset-lukevat-aiempaa-vahemman-kirjoja-netissa-ja-somessa-lukutaito-maarittyy-uudelleen/>

- 109 Kauppinen, M., ja Marjanen, J. (2020). Millaista on yhdeksäsluokkalaisten kielellinen osaaminen? Suomen kielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa 2019. Kansallinen koulutuksen arviointikeskus. https://www.karvi.fi/sites/default/files/sites/default/files/documents/KARVI_1320.pdf
- 110 TrustLab, (2023). Code of Practice on Disinformation, Semi-annual Report, September 2023. <https://disinfocode.eu/wp-content/uploads/2023/09/code-of-practice-on-disinformation-september-22-2023.pdf>
- 111 Kiili, K., Hirvonen, H., Hämäläinen, E., Kanninen, L., Kiili, C., Kujju, P., Kupiainen, R., Muukkonen, P., Paakkari, L., Ruotsalainen, J., ja & Torppa, M. (2023). Nuoret tarvitsevat enemmän tukea kriittiseen lukemiseen. https://educritical.fi/wp-content/uploads/2023/03/politiikkasuositus_CRITICAL_27_03_2023.pdf
- 112 Kauppinen, M., ja Marjanen, J. (2020). Millaista on yhdeksäsluokkalaisten kielellinen osaaminen? Suomen kielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa 2019. Kansallinen koulutuksen arviointikeskus. https://www.karvi.fi/sites/default/files/sites/default/files/documents/KARVI_1320.pdf
- 113 Opetushallitus, (2016). Perusopetuksen opetussuunnitelman perusteet 2014. https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf
- 114 Pönkä, Harto (2016). Monilukutaito ja sosiaalinen media opetuksessa. Teoksessa Kaisa Leino ja Outi Kallionpää (toim.), Monilukutaitoa digiaikaan – lukemisen ja kirjoittamisen uudet haasteet ja mahdollisuudet, Äidinkielen opettajain liiton vuosikirja 2016. <https://harto.wordpress.com/2016/03/30/monilukutaito-ja-sosiaalinen-media-opetuksessa-artikkeli/>



9. Mitä tarkempi syöte – sen parempi tulos

KARI KIVINEN

GenAI:n toiminta perustuu sille annettuihin syötteisiin. Saatua tulos riippuu paljon siitä, miten huolella syöte tai kysymys on laadittu. Syöte tekstien taitavasta suunnittelusta (engl. prompt engineering) on tullut tärkeä taito generatiivisten tekoälyjärjestelmien tehokkaan

käytön ja käyttöönoton kannalta. Syötteessä olisi hyvä kertoa, mikä rooli tekoälyllä on, mitä haluaa sen tekävän, antaa selkeitä ohjeita siitä, miten edetä sekä tarpeeksi taustatietoa tehtävän suorittamiseksi.

Ohjeita syötteiden käytöstä¹⁵

Syötettä tai kysymystä on usein tarpeellista hioa useita kertoja. Tekoälyn tuotokset saattavat olla arvaamattomia. Jos syöte ei toimi, muuta sitä ja yritä uudelleen.

Olet vastuussa omasta työstäsi. Vaikka tekoäly voi luoda hetkessä vaikuttavan oloisen tuotoksen, niin sen sisältö voi olla virheellinen. Lopullinen työ kannattaa tarkistaa luotettavista lähteistä.

Tekoäly voi antaa realistisia vääriä vastauksia: Koskaan ei pidä hyväksyä sen tuotosta sellaisenaan. Ole erityisen varovainen lähteiden, faktojen tai lainausten suhteen, sillä ne ovat hyvin todennäköisesti virheellisiä.

Lisensioituja aineistoja, esim. tieteellisiä artikkeleita, ei saa syöttää mihinkään kielimalleihin tai muihinkaan tekoälytyökaluihin. Artikkelin CC-lisenssit eivät kumoja tätä sääntöä

Älä jaa tekoälyn kanssa mitään henkilökohtaista. Kaikkea jakamaasi tietoa voidaan käyttää tekoälyn harjoitustietona.

Tyydyttävän tuloksen saamiseksi on välillä nähtävä vaivaa. Syötteitä kannattaa kokeilla, kehittää ja muuttella. Samalla kannattaa tarkkailla, miten se vaikuttaa vastauksen sisältöön ja laatuun. Oheisessa esimerkissä ChatGPT:tä pyydetään muuntamaan Wikipediasta ko-

pioitua tekstiä suomea vieraana kielenä opiskelevien oppilaiden käyttöön sopivaksi. Kokemus on osoittanut, että ChatGPT toimii melko luottavasti, kun sille on syötetty omaa tekstiaineistoa.

Esimerkki

Syöte	Muunna oheinen teksti sellaiseksi selkokieliseksi tekstiksi, jonka peruskoulun 2. luokan oppilas, joka opiskelee suomea S2-kielenä, voisi ymmärtää helposti. Älä tee mitään muuta.
Tekstisyöte (Lähde Wikipedia)	<p>Vesinokkaeläin (<i>Ornithorhynchus anatinus</i>) on pienikokoinen nisäkäslaji, jota esiintyy Kaakkois-Australiassa ja Tasmaniassa. Neljän nokkasiililajin ohella vesinokkaeläin on yksi viidestä nokkaeläimiin kuuluvasta nisäkäslajista, jotka eivät synnytä eläviä poikasia vaan munivat. Muiden nokkaeläinten tapaan myös vesinokkaeläimellä on sähköaisti.</p> <p>Muniva, myrkyllinen, räpyläjalkainen ja ulkomuodoltaan ankannokkaiseksi sekä majavahäntäiseksi kuvailtu nisäkä on hämmäntänyt eurooppalaisia eläintieteilijöitä, joista monet ensi kohtaamisella luulivat sitä hyvin suunnitelluksi huijaukseksi. Se on yksi harvoista myrkyllisistä nisäkkäistä: vesinokkaeläinuroksella on takajalassaan kannus, joka levittää ihmiselle kovaa kipua aiheuttavaa myrkyä. Vesinokkaeläin on ainutlaatuisten ominaisuuksiensa vuoksi tärkeä evoluutiobiologian tutkimuskohde ja Australian helposti tunnistettava symboli. Sitä on käytetty useiden kansallisten tapahtumien maskottina, ja se on kuvattuna Australian 20 sentin kolikon kääntöpuolelle.</p>
ChatGPT:n muuntama teksti	<p>Vesinokkaeläin on pieni eläin, joka asuu Australiassa. Se on yksi viidestä erityisestä eläimestä, jotka kuuluvat nokkaeläimiin. Nämä eläimet eivät synnytä vauvoja, vaan ne munivat munia.</p> <p>Vesinokkaeläin näyttää vähän oudolta. Se on pieni, sillä on nokka kuin ankalla ja pyrstö kuin majavalla. Se on myös myrkyllinen, eli se voi tehdä kipeää, jos se iskee sinua takajalkansa kannuksella. Vesinokkaeläimellä on erityinen aisti, joka auttaa sitä löytämään saaliinsa.</p> <p>Kun ensimmäiset eurooppalaiset näkivät vesinokkaeläimen, he eivät uskoneet, että se on oikea eläin. He ajattelivat, että se on joku pilailu! Mutta vesinokkaeläin on todellinen ja se on tärkeä eläintutkimuksen kohde. Se on myös Australian symboli, ja voit nähdä sen kuvan Australian 20 sentin kolikossa.</p>

Kuvantuottamisessa kannattaa määrittellä todella tarkasti, minkälaista lopputulosta tavoittelee. Mitä selkeämmin ja monipuolisemmin kuvailtu syöte, sen parempi tulos.

- Anna mahdollisimman selkeä selite kuvasta, jonka haluat luoda. Ole mahdollisimman tarkka, jotta tekoölygeneraattori tietää, mitä etsiä tietokannoissaan.
- Lisää syötteeseen haluamiasi yksityiskohtia; kuvan koko, muoto, tekstuurit ja mahdolliset värit tai sävyt.
- Kuvaile tarkasti millaisen esineen/ihmisen/eläimen haluat kuvaan ja minkälaisessa ympäristössä se on.
- Kerro, haluatko tehdä piirroksen, akvarellin, öl-

jyvärityön, hiilipiirroksen vai kenties 3D-kuvan? Tavoitteletko valokuvamaista lopputulosta, graffita, animea tai jotain muuta.

- Määritä kuvan tyyli tai tyyllilaji. Voit käyttää esim. avainsanoja "realistinen", "valokuvamainen", "abstrakti", "minimalistinen", "kubistinen", "surrealistinen" jne. Joissain ohjelmissa voidaan antaa myös vihje taitelijasta esim. Dalin tai Picasson tapaan".
- Voit vaikuttaa myös esimerkiksi kuvan sommiteluun, valaistukseen, kuvasuhteeseen, kamerakulmaan, resoluutioon, valaistustyyliin ja kameranäkymään.
- Kokeile erilaisia syötteitä, muokkaa niitä ja leiki rohkeasti sovelluksen kanssa!

Esimerkki kuvasyötteestä

Tee värikäs akvarellimainen piirros papukaijasta kukikasta verhotaustaa vasten. Punavihreä papukaija istuu tangolla ja se on sanomassa jotakin. Lisää puhekupla. Kuvaa papukaijaa sivulta hyvässä valaistuksessa.



(www.bing.com/create)

¹¹⁵ Mollick, E.& Mollick, L. (2023) Assigning AI: seven approaches for students, with prompts. <http://dx.doi.org/10.2139/ssrn.4475995>

10. Tekoälykäästä oppimista?

KARI KIVINEN

Algoritmeista, tekoälystä ja generatiivisesta tekoälystä on tullut osa jokapäiväistä arkeamme. Niitä käytetään yleisesti jo erilaisissa sovelluksissa kuten esimerkiksi hakukoneissa, chattiboteissa, käännöskoneissa ja navigointisovelluksissa. ChatGPT:n julkistamisen jälkeen markkinoille on ilmestynyt lukemattomia sovelluksia, jotka ovat avanneet tekoälyn hyödyntämismahdollisuuksia aivan uusille aloille – myös koulutukseen. Niiden käyttö on tehty mahdollisimman helpoksi ja vaivattomaksi. Jokainen, joka on käyttänyt hakukoneita, osaa myös tehdä ChatGPT-syötteitä (engl. prompts). Sovellukset toimivat usein monilla kielillä ja myös suomeksi. Generatiiviset tekoälyohjelmat ovat osoittautuneet tehokkaiksi työkaluiksi, joiden käytössä on kuitenkin omat niksinsä ja haasteensa.

Otavia on julkaissut opettajille suunnatun Tekoäly tutuksi-oppaan, jonka avulla on helppo tutustua ChatGPT:n käyttöön.¹¹⁶

Etäopetuskauden myötä suurin osa eurooppalaisista kouluista siirtyi käyttämään erilaisia digitaalisia alustoja¹¹⁷ ja interaktiivisuutta elävöittäviä ohjelmistoja. Siksi siirtyminen tekoälyn käyttäjiksi on tapahtunut verrat- tain luontevasti. Koulutusjärjestelmien digitalisointi on lisännyt¹¹⁸ tallennettavan, kopioitavan, siirrettävän ja käsiteltävän datan määrää, eivätkä käyttäjät aina tiedä, missä ja miten heidän tietojensa käytetään. Emme esimerkiksi tiedä tarkasti, mitä tietoa ja dataa digitaaliset alustat ovat keränneet oppilaista, opettajista ja opetussisällöistä viime vuosien aikana. Esimerkiksi monien käyttämä Zoom joutui kesällä 2023 vakuuttamaan käyttäjilleen, ettei se käytä ääni-, video-, chat-, ruudun- joko-, liitetiedostoja tai asiakassivaltöä (kuten mielipide- kyselyjen tuloksia, valkotauluja ja reaktioita) Zoomin tai kolmannen osapuolen tekoälymallien kouluttamiseen ilman käyttäjän suostumusta.¹¹⁹

Generatiivisen tekoälyn käytön räjähdysmäinen li-

tekevät siitä arvokkaan välineen esimerkiksi asiakaspalvelun ja kielenkääntämisen kaltaisiin tehtäviin. Nämä työkalut ovat kuitenkin vain yhtä tehokkaita kuin niiden saamien inhimillisten kehotusten laatu.

Tekoälyn kehityskaaren hahmottaminen: Tekoälylukutaito edellyttää myös tekoälyn historian ja tulevaisuuden kehityksen ymmärtämistä. Kehittyneiden tekoälyteknologioiden, kuten GPT-4:n kaltaisten suurten kielimallien, käyttöönotto on aiheuttanut samankaltaisia mullistavia muutoksia kuin veturin vaikutus teollisen vallankumouksen aikana. Tämä kehitys kuvastaa siirtymistä keinotekoisesta suppeasta älykkyydestä (Artificial Narrow Intelligence, ANI) keinotekoiseen yleiseen älykkyyteen (Artificial General Intelligence, AGI), mikä lupaa merkittäviä tuottavuushyppäyksiä tulevina vuosina.¹³⁰ Tulevat muutokset voivat muokata yhteiskuntaa ja henkilökohtaisia vuorovaikutussuhteita uudelleen mittakaavassa, joka ylittää aiemmat teknologiset edistysaskeleet. ChatGPT:n edistyminen kohti AGI:tä on kiistatonta, ja näiden teknologioiden kehittyessä ne tarjoavat sekä haasteita että mahdollisuuksia parantaa ja laajentaa edelleen eri aloille). Näiden suuntausten tunnistaminen ja ennakoiminen mahdollistaa tämän tekoälyvallankumouksen mahdollisten hyötyjen optimaalisen hyödyntämisen.

Yleisesti ottaen tekoälylukutaitoa pidetään 2000-luvulla välttämättömänä taitona kaikille, ei vain tietojenkäsittelytieteilijöille.¹³¹

Tekoäly on tullut jäädäkseen opetuksen avuksi

Kehittyneet tekoälysovellukset tarjoavat erinomaisia työkaluja esimerkiksi opetuksen ja koulun hallintorutiinien kehittämiseen. Tekoälypohjaisia oppimista avustavia ohjelmistoja ja sovelluksia on jo jokaisen koulun käytettävissä.

ChatGPT:n julkaisemisen ensireaktiot keskittyivät suurelta osin uuden teknologian mahdollisiin väärinkäytöksiin ja ”tekoälyn avulla huijaamiseen”.¹³² Syksyllä 2023 painopiste on siirtymässä tapoihin, joilla generatiivinen tekoäly voisi tukea oppimista ja opetusta.¹³³ Monissa ehdotetuissa opetussovelluksissa ChatGPT tekee yhteistyötä oppijan kanssa, esimerkiksi vastaväittäjänä, yhteissuunnittelijana, motivaattorina, tutorina tai opiskelukumppanina. Taidekoulutusta koskevassa artikkelissa Nate Harrison toteaa: ”Emme voi taistella tekoälyä vastaan. Opiskelijat käyttävät sitä jo nyt, monet hyvin organisaatiosta, intuitiivisesti ja luonnollisesti.” Hänen mielestään kysymys kuuluu: ”Miten voimme varustaa opiskelijat älykkäillä lähestymistavoilla tekoälyn käyttöön?”¹³⁴

ChatGPT:n kaltaiset tekoälyjärjestelmät olisivat hyvä nähdä välineinä, joilla tuetaan ja täydennetään opetusprosessia, ei sen korvikkeena. Tutkijat pyysivät ChatGPT:tä antamaan vastauksen kysymykseen, mihin sitä voidaan käyttää opetuksessa. Samalla he arvioivat kriittisesti sinänsä hyvältä vaikuttavia, mutta tekoälylle tyypilliseen tapaan ympäröiväitä vastauksia (taulukossa ChatGPT:n ehdotus ja refererointi tutkijoiden vastineesta):

Mihin ChatGPT:tä voi käyttää opetuksessa?¹³⁵

1	Kysymyksiin vastaaminen: ChatGPT:n avulla voidaan vastata kysymyksiin reaaliaikaisesti, jolloin opiskelijat saavat välitöntä palautetta ja voivat oppia tehokkaammin.
Vastine	ChatGPT:n käyttäminen käsitteiden selventämiseen opiskelijoille edellyttää, että opiskelijoilla on taito kirjoittaa syötteitä: vain esittämällä oikeita kysymyksiä saadaan päteviä vastauksia. Lisäksi, koska opiskelijat ovat aloittelijoita omalla alallaan, heidän on vaikea määrittää, hallusinoiko ChatGPT vai ei.
2	Tutorointi: ChatGPT voidaan integroida opetusohjelmistoihin, jotta oppilaat voivat saada henkilökohtaista opetusta.
Vastine	Se, että keskustelu ChatGPT:n kanssa vaikuttaa lähes inhimilliseltä ja on usein kohteliasta ja ystävällistä, sisältää riskin, että käyttäjä ei ole tietoinen siitä, että hän on yhteydessä tekoälyjärjestelmään, jonka toiminnallisuutta ja tarkkaa harjoitusdataa ei tunneta.

3	Sisällön tuottaminen: ChatGPT:tä voidaan käyttää tuottamaan opetussisältöä, kuten yhteenvetoja, selityksiä ja oppimateriaalia, saatujen tietojen perusteella.
Vastine	ChatGPT tarjoaa esimerkkejä opetussisällön tuottamiseen (tiivistelmät, selitykset ja oppimateriaalit). Sitä voidaan käyttää myös tekstin tuottamiseen opiskelijoiden tehtäviin, kuten kirjallisuuskatsauksiin. Tällöin on olennaista, että ihminen tarkistaa sisällön, jotta voidaan varmistaa, ettei opetusmateriaali ole saastunut ennakkoluuloista tai vääränetystä tiedosta.
4	Kielten oppiminen: ChatGPT:tä voidaan käyttää kielen oppimisen ohjaajana, joka auttaa opiskelijoita harjoittelemaan puhumista, kirjoittamista ja ymmärtämistä vieraalla kielellä.
Vastine	ChatGPT ei harjoita kriittistä itsereflektiota kielenoppimisen kannalta. Koska se ei (vielä) kykene ”ymmärtämään” vivahteita puheessa eikä ottamaan huomioon kulttuurieroja, on kysyttävä, voiko se todella tukea puhumista, kirjoittamista ja ymmärtämistä vieraalla kielellä. Kun tarkasteltiin ChatGPT:n käyttöä kieltenopetuksessa, se näytti edelleen kampailevan oikeiden käänntösten tuottamisessa kielillä, joita oli harjoittelukokonaisuudessa niukasti. Nämä vaikeudet on kuitenkin ehkä jo ratkaistu seuraavissa versioissa.

ChatGPT:n tulosten analysoinnin lopputuloksena tutkijat tulivat siihen johtopäätökseen, että ”digitaalisen lukutaidon ja kriittisen ajattelun kehittäminen edellyttää inhimillisiä opettajia. Vasta kun oppilaat ovat kehittäneet näitä taitoja ja ymmärtävät, miten ChatGPT:n tuottaman sisällön arvoa voidaan verrata heidän oman kirjallisen tiedonkäsittelynsä kautta tuotettuun sisältöön, ChatGPT voi parantaa heidän oppimistaan.” Vastaavasti opettajat hyötyvät tekoälystä osana opetusta, kun he hahmottavat, miten ChatGPT:tä käytetään viisaasti.

Tekoälyjärjestelmiä voidaan käyttää esimerkiksi erityisten oppimistarpeiden tunnistamiseen ja opetuksen eriyttämiseen. Kouluhallinnon ohjelmistoilla voidaan optimoida esimerkiksi lukujärjestyksiä ja tehostaa tilojen ja resurssien käyttöä. Monikulttuurisessa luokassa tekoälyä voidaan käyttää käänntösten tekemiseen ja opetusmateriaalin eriyttämiseen. ChatGPT:n avulla voi muuntaa oppimateriaalin vaikkapa S2-oppilaiden kielitasoa vastaavaksi, kuten artikkelin lopun esimerkistä käy ilmi.

Monet opettajat ovat pelänneet tekoälyn vähentävän opettajien tarvetta. Tällä hetkellä näyttää siltä, että huoli on turha. Vaikka tekoälypohjaiset opetusohjelmat kehittyvät nopeasti, on ilmeistä, ettei tekoäly tule koskaan korvaamaan inhimillistä sosiaalista vuorovaikutusta – opettajan ja oppilaan kohtaamista. Oppiminen on vuorovaikutteinen prosessi, jossa oppija rakentaa tietämystään aikaisempien tietojen pohjalta vuorovaikutuksessa toisten kanssa. Tekoälypohjaiset sovellukset voivat opettajan johdolla toimia hyvinä avustajina.

Tekoäly voi tukea opettajan työtä ja sen avulla voidaan esim. automatisoida hallinnollisia rutiinitehtäviä. Opettajat voivat nykyään käyttää tekoälyä apuna esimerkiksi oppilastöiden arvioinnissa. Tuoreen tutki-

muksen mukaan¹³⁷ ChatGPT pystyy arvioimaan opiskelijoiden tekstejä melkein yhtä hyvin kuin opettajat. Jos otetaan huomioon palautteen antamisen helppous ChatGPT:n avulla ja sen tuottamien tekstien laatu, erot ihmisen ja ChatGPT:n välillä eivät ole merkittäviä. Onkin helppo ennustaa, että oppilastuotosten arviointi siirtyä lähitulevaisuudessa tekoälyavusteiseksi. Silloin olisi syytä pohtia kriittisesti oppilaiden oikeusturva: tekeekö tekoälyohjelmisto oikeudenmukaisia päätöksiä ja millaisia mahdollisuuksia opettajilla on tarkkailla ja tarkistaa automaattisesti tehtyjä arviointeja ja päätöksiä?

Tekoälyn käyttöönotto

Opettajien ja koulujen johdon olisi voitava osallistua tekoälysovellusten kehittämiseen ja edistää siten opetusta tukevien tekoälyjärjestelmien eettistä, luotettavaa, oikeudenmukaista ja osallistavaa kehittämistä ja käyttöä.

Komission asiantuntijaryhmä, joka pohti tekoälyn ja datan hyödyntämistä opetuksessa, ehdotti monenlaisia toimenpiteitä tekoälyjärjestelmien käyttöönotossa kouluissa.¹³⁸

Lisäksi koulutuksen järjestäjien olisi syytä varmistaa, että tekoälyjärjestelmät ovat tekoälyä koskevan EU:n tekoälysäädöskehityksen mukaisia. Heidän tulisi keskittyä siihen, miten opettajat ja oppijat voisivat käyttää tekoälyä ja dataa eettisellä ja tietosuojasääntösten mukaisella tavalla opetuksessa, oppimisessa ja arvioinnissa.

Kun tekoälyjärjestelmät ja niitä koskeva säännöstö kehittyvät nopeasti ja datan käyttö lisääntyy, koulutuksen järjestäjien tulisi huolehtia siitä, että hallinto- ja opetushenkilöstö pysyy ajan tasalla ja saa täydennyskoulutusta.

Mitä tekoälyjärjestelmien hankinnassa ja käyttöönotossa julkisella sektorilla olisi huomioitava?

(Muokattu ja täydennetty suosituslista.)

- Varmistetaan, että julkisilla hankinnoilla hankitaan luotettavia ja ihmiskeskeisiä tekoälyjärjestelmiä.
- Noudatetaan yleisen tietosuoja-asetuksen velvoitteita ja toteutetaan tietosuojaa koskevaa jatkuvaa arviointia.
- Toteutetaan ihmisen suorittamaa valvontaa.
- Tarjotaan henkilöstölle ja käyttäjille asianmukaista koulutusta.
- Seurataan tekoälyjärjestelmän toimintaa ja toteutetaan korjaavia toimia tarpeen mukaan.
- Suhtaudutaan terveeseen kriittisesti tekoälyn tuotoksiin.
- Ohjeistetaan henkilökuntaa, ettei tekoälysyötteisiin sisällytetä mitään luottamuksellista tietoa.
- Ohjeistetaan, ettei tekoälyn tuotoksia jaeta ilman ihmisen suorittamaa tarkistusta.

Asiantuntijat tuntevat olevan Euroopan komission tutkimuskeskuksen tuoreen raportin mukaan yhtä mieltä siitä, että generatiivisella tekoälyllä on pysyvä vaikutus – sekä myönteinen että kielteinen – opetukseen, oppimiseen ja oppimiseen ja arviointiin.¹³⁹ – Generatiivisen tekoälyn uusimmat sovellukset tarjoavat monia uusia mahdollisuuksia pedagogisiin innovaatioihin. Opetuksessa voidaan hyödyntää esimerkiksi automaattista videotekstitystä, tekstin muuntamista puheeksi tai puheen muuntamista tekstiksi, kääntämistä kieleltä toiselle sekä esitysten, animaatioiden ja kuvien luomista tekstin pohjalta.

Kun tekoälyjärjestelmistä kehittyä oppimista tukevia ”älykkäitä kumppaneita”, ”oppimiskumppaneita” tai ”kognitiivisia työkaluja”, niin metakognitio ja pohdinta nousevat tärkeämmiksi kuin ulkoa opettelu. Tämä merkitsee toimijuuden uudelleenjakoa ihmisten ja koneiden välillä. Tällöin on tärkeää tehdä tietoon perustuvia päätöksiä siitä, mitä toimintoja voimme ”delegoida” tekoälyavusteisille järjestelmille ja mitä toimintoja meidän tulisi säilyttää itsellämme.

Komission raportissa tunnustetaan myös tekoälyn riskit. On seurattava ja arvioitava kriittisesti esimerkiksi kaupallisten etujen korostumisen riskejä tai digitaalisen kuilun syvenemistä. Sillä tarkoitetaan digitaalitekniikan eriarvoista saatavuutta.

ChatGPT:n kehittäjä OpenAI julkaisee blogeja opettajille, jotka käyttävät ChatGPT:tä opetuksessaan. Ne sisältävät lisätietoa ChatGPT:n toiminnasta, sen rajoituksista, tekoälyilmaisimien tehokkuudesta (joiden avulla voi tarkistaa, onko teksti tuotettu tekoälyllä) sekä tuotosten vinoumista.¹⁴⁰

Luotettavan tekoäly-ympäristön eettiset vaatimukset

Komission oppaassa ”Tekoälyn ja datan käyttö opetuksessa – eettiset ohjeet opettajille” esitetään luotettavaa tekoälyä koskevat keskeiset eettiset vaatimukset:

1. **Ihmisen toimijuus ja ihmisen suorittama valvonta** – sisältää esimerkiksi perusoikeuksien ja lasten oikeuksien turvaamisen, ihmisen toimijuiden ja ihmisen suorittaman valvonnan.
2. **Läpinäkyvyys** – sisältää esimerkiksi jäljitettävyyden, selitettävyyden ja tiedotuksen.
3. **Monimuotoisuus, syrjimättömyys ja oikeudenmukaisuus** – sisältää esimerkiksi esteettömyyden, kaikkia palvelevan suunnittelun, epäoikeudenmukaisten vääristymien välttämisen ja sidosryhmien osallistumisen. Periaatteet mahdollistavat järjestelmien käytön iästä, sukupuolesta, kyvyistä tai ominaisuuksista riippumatta kiinnittäen erityistä huomiota opiskelijoihin, joilla on erityistarpeita.
4. **Yhteiskunnallinen ja ekologinen hyvinvointi** – sisältää esimerkiksi kestävyuden ja ympäristöystävällisyyden, sosiaaliset vaikutukset, yhteiskunnan ja demokratian.
5. **Yksityisyyden suoja ja datan hallinta** – sisältää esimerkiksi yksityisyyden kunnioittamisen sekä datan laadun, eheyden ja saatavuuden.
6. **Tekninen luotettavuus ja turvallisuus** – sisältää esimerkiksi vastustuskyvyn kyberhyökkäyksiä vastaan, turvallisuuden, tarkkuuden, luotettavuuden

den ja toistettavuuden.

7. **Vastuuvollisuus** – sisältää esimerkiksi tarkastettavuuden, kielteisten vaikutusten minimoinnin ja niistä ilmoittamisen, valinnat ja muutoksenhaun.

Näiden perusvaatimusten tarkoituksena on edistää luotettavien tekoälyjärjestelmien kehittämistä ja käyttöönottoa. Ne voivat auttaa rehtoreita ja opettajia hyödyntämään tekoälyn käyttöä, arvioimaan sen käytön vaikutuksia ja puuttumaan mahdollisiin riskeihin. (Ks. Liite 2: Komission asiantuntijaryhmän ohjaavat kysymykset opettajille.)

Hybridioppimisen hyödyt

Hybridioppimisella tarkoitetaan prosessia, jossa tekoäly ja ihmiset toimivat yhdessä edistääkseen oppimisen kognitiivisia ja metakognitiivisia puolia.

Professori Sanna Järvelän mukaan lasten ja nuorten kyky säädellä omaa oppimistaan ja siihen liittyviä taitoja on kriittisen tärkeää tekoälyn aikakaudella.¹⁴² Tekoälyohjattujen oppimistekniikoiden käyttö yleistyy kouluissa eri puolilla maailmaa. Koska laajamittaista tutkimusta ja käytettävissä olevia ratkaisuja ei vielä ole, tekoälyn mahdollisuuksia ei tunneta tarpeeksi. Kouluissa tarvitaan teknologioiden ja pedagogiikan yhdistämistä. Oulun yliopisto on mukana Hybrid Intelligence (HI) tutkimusohjelmassa.¹⁴³

Tekoälyä voi esimerkiksi käyttää tutorina tai valmentajana oppimisprosessin säätelyssä, jolloin oppijan ja teknologian välille syntyy todellinen hybridisuhte. Oppija voi pyytää generatiiviselta tekoälyltä apua esimerkiksi tiedonkäsittelyssä, ongelmanratkaisussa tai sisällön luomisessa, jolloin osa kognitiivisista tehtävistä jää pois ja oppija voi keskittyä korkeamman asteen ajattelutaitoihin. Generatiivisen tekoälyn vastaukset edellyttävät kuitenkin vastausten arviointia, niiden merkityksen selvittämistä ja mahdollisesti lisäkysymysten esittämistä, koska generatiivisen tekoälyn tarjoaman tiedon luotettavuutta ja tarkkuutta ei voida taata. Työskentely generatiivisen tekoälyn kanssa edellyttää oppijalta kehittyneitä itsesäätelytaitoja. Reaaliaikaisen palautteen, mukautuvan kysymyksenasettelun ja itsearviointikehotteiden kaltaisten ominaisuuksien avulla tekoäly voi ohjata oppijoita seuraamaan, arvioimaan ja mukautamaan oppimisstrategioitaan ja siten parantamaan metakognitiivisia taitojaan.¹⁴⁴

EdGPT: tekoälymalli koulutusta varten

Unescon raportissa hahmotellaan generatiivisen tekoälyn ja koulutuksen lähitulevaisuutta. Tekoälyn taustatilat ovat alkaneet kehittää ”EdGPT:tä”, jota koulutetaan palvelemaan opetustarkoituksia. Toisin sanoen

EdGPT:n kehittäjien tavoitteena on luoda malli, joka on johdettu valtavista määristä yleistä harjoitusdataa ja pienemmillä määrillä korkealaatuista, alakohtaista koulutusdataa.¹⁴⁵

Tällä hetkellä GPT-mallien kohdentaminen koulutukseen on kuitenkin vasta alkuvaiheessa. Esimerkiksi TAL Education Group on kehittänyt MathGPT:n, joka keskittyy matematiikkaan liittyvään ongelmanratkaisuun maailmanlaajuisesti.¹⁴⁶ Haasteena koulutusprojekteilla on määritellä, missä määrin oppija- ja opettajatietoja voidaan eettisesti kerätä ja käyttää EdGPT:n perustana.

Teoriassa EdGPT-mallien tuotokset voisivat myös sisältää vähemmän yleisiä virheitä tai muuta epäsuotuisaa sisältöä kuin tavallinen GPT. Ne saattavat silti edelleen tuottaa virheitä, ellei taustalla olevia GenAI-malleja ja -lähestymistapaa muuteta merkittävästi.

Suosituksia tekoälyn vastuulliseen hyödyntämiseen

Digitaalisia oikeuksia ja periaatteita koskevassa eurooppalaisessa julistuksessa sanotaan: ”Jokaisella on oikeus koulutukseen ja elinikäiseen oppimiseen, ja jokaisella olisi oltava mahdollisuus hankkia kaikki digitaaliset perustaidot ja edistynyt digiosaamista.”¹⁴⁷

Euroopan tekoälystrategian tavoitteena on tehdä EU:sta maailmanluokan tekoälykeskus ja varmistaa, että tekoäly on ihmiskeskeistä ja luotettavaa. Komission mukaan tapa, jolla suhtaudumme tekoälyyn, määrittelee maailman, jossa elämme tulevaisuudessa. EU:n kansalaisten ja yritysten olisi voitava nauttia tekoälyn hyödyistä ja samalla tuntea olonsa turvallisiksi ja suojatuksi. Osana digitaalistrategiaansa Euroopan komissio on ehdottanut tekoälyä koskevaa kattavaa oikeudellista kehystä (tekoälysäädös¹⁴⁸), jota tullaan soveltamaan monilla aloilla – myös yleissivistävässä ja ammatillisessa koulutuksessa. Säädös perustuu tekoälyä ja dataa koskevan EU:n säätelyn ja politiikan kehitykseen.

Digitaalisen koulutuksen toimintasuunnitelma (2021–2027) on EU:n poliittinen aloite, jossa esitetään visio laadukkaasta, osallistavasta ja helppokäyttöisestä digitaalisesta koulutuksesta Euroopassa ja jolla pyritään tukemaan jäsenvaltioiden koulutusjärjestelmien mukautamista digitaali-aikaan.¹⁴⁹

Tässä artikkelissa on viitattu useaan otteeseen ”Tekoälyn ja datan käyttö opetuksessa ja oppimisessa – eettiset ohjeet opettajille”¹⁵⁰ julkaisun suosituksiin. Digitaalisessa informaatiolukutaito-oppaassa¹⁵¹ annetaan lisätietoja kansalaisten perusdigitaaloista (DigComp 2.2.¹⁵²). Sen AI-liitteen tekoälyä koskevista osaamiskuvuksista kerrotaan lisää tämän oppaan luvussa Tekoäly – mitä meidän pitäisi tietää ja osata?

Lisäksi suomeksi on käännetty myös Euroopan komission tuottaman toimintasuunnitelman kolmas osa: ”Ohjeet opettajille ja kasvattajille disinformaation tor-

jumisesta ja digilukutaidon edistämisestä koulutuksen avulla”.¹⁵³

Suomessa Opetushallitus on järjestänyt syksyllä 2023 pyöreän pöydän keskustelun tekoälyn etiikasta opetuksessa. Näitä keskusteluita on tarkoitus jatkaa.

Suomalaisten yliopistojen Generation AI-yhteishanke yhdistää tutkijat, teknologian kehittäjät, koulut, viranomaiset, yritykset ja kansalaisjärjestöt.¹⁵⁴ Hankkeen tavoitteena on vahvistaa lasten ja nuorten kykyä kohdata teknologian globaaleja yhteiskunnallisia vaikutuksia ja tuottaa opettajien käyttöön pedagogisia työkaluja oppimisen murrokseen. Hankkeessa rakennetaan perustaa, jonka päälle voidaan rakentaa ”AI-sukupolven” teknologiakasvatusta, joka ottaa huomioon tekoälyn toimintaperiaatteet, sen tarjoamat mahdollisuudet ja dynaamiset vaikutukset – unohtamatta sen puutteita ja riskejä tai sitä kuinka tekoälyjärjestelmät voivat kasvat-
taa vastakkainasettelua, syrjintää ja turvattomuutta tai rapauttaa luottamusta.

Minun tekoälyni

Generatiivisen tekoälyn kannattaa tutustua testamalla sen käyttöä kaikessa rauhassa itsekseen ja ko-keilemalla rohkeasti erilaisten syötteiden käyttöä ja tar-
kentamalla niitä lisämäärittelyillä. Kuvageneraattoreiden kanssa leikkiminen havainnollistaa hienosti syötteiden tarkkuuden merkityksen – sitä saa mitä tilaa!

Sosiaalisessa mediassa on jo useita sivustoja¹⁵⁵, joi-
sa opettajat jakavat kokemuksiaan ja hyviä käytänteitä toistensa kanssa ja monet tahot järjestävät opettajille suunnattuja tekoälyn käyttöä tukevia koulutuksia¹⁵⁶ ja tapahtumia¹⁵⁷.

Meneillään on taas kerran teknologian murroskausi. Uudet tekoälysovellukset muuttavat elämäämme, halu-
simme sitä tai emme. Siksi on parempi ottaa asioista selvää ja hyödyntää työssä ja vapaa-ajalla tekoälyn tuomia uusia mahdollisuuksia tiedostaen kuitenkin sen puutteet ja vajavaisuudet.

¹¹⁶ Otavia (2023) Tekoäly tutuksi. https://sites.google.com/otavanopisto.fi/tekoaluyoppimisen_tukena?fbclid=IwAR3wReAltN2KTmyb3kY2lTiEX9AWu3lOM7ELW4LJMvYyI5s5jb-jllZpZf8

¹¹⁷ Eurydice (2023) Teaching and learning in schools in Europe during the Covid-19 pandemic. <https://eurydice.eacea.ec.europa.eu/publications/teaching-and-learning-schools-europe-during-covid-19-pandemic>

¹¹⁸ European Commission, Directorate-General for Education, Youth, Sport and Culture, Final report of the Commission expert group on artificial intelligence and data in education and training – A executive summary, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/65087>

¹¹⁹ Zoom blog (2023) How Zoom's terms of service and practices apply for AI features. <https://blog.zoom.us/zooms-term-service-ai/>

¹²⁰ UNESCO survey (2023) Less than 10% of the schools and universities have formal guidance on AI. <https://www.unesco.org/en/articles/unesco-survey-less-10-schools-and-universities-have-formal-guidance-ai>

¹²¹ Fengchun, M. & al (2023) Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

¹²² European Commission, Directorate-General for Education, Youth, Sport and Culture, Final report of the Commission expert group on artificial intelligence and data in education and training – A executive summary, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/65087>

¹²³ Hirvonen, N. & Enwald, H. 2022. Tekoälyn lukutaito – valmiuksia, joita jo tarvitsemme. Signum, 53(4), 25–30.

¹²⁴ Matthias Carl Laupichler, Alexandra Aster, Jana Schirch, Tobias Raupach. 2022. Artificial intelligence literacy in higher and adult education: A scoping literature review, Computers and Education: Artificial Intelligence, Volume 3, <https://doi.org/10.1016/j.caeai.2022.100101>

¹²⁵ Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. Paper presented at the Proceedings of the 2020 CHI conference on human factors in computing systems

¹²⁶ Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., . . . Gašević, D. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? Computers and Education: Artificial Intelligence, 3, 100056.

¹²⁷ Viriya Techarungroj & Kanok Karnchanapoo (2023) AI Literacy in the Age of Generative AI – Building a Future-Ready Society AI Governance Clinic Expert Fellows. EDTA. <https://www.edta.or.th/getattachment/Our-Service/ALGC/Research-and-Recommendation/AI-Literacy-Viriya-and-Kanok.pdf.aspx?lang=th-TH>

¹²⁸ Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., . . . Ahuja, M. (2023). “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642.

¹²⁹ Gašević, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. Computers and Education: Artificial Intelligence, 100130.

¹³⁰ Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., . . . Ahuja, M. (2023). “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642.

¹³¹ Davy Tsz Kit Ng, Jac Ka Lok Leung, Samuel Kai Wah Chu, Maggie Shen Qiao. 2021. Conceptualizing AI literacy: An exploratory review, Computers and Education: Artificial Intelligence <https://doi.org/10.1016/j.caeai.2021.100041>

¹³² Tuomi, I., Cachia, R. and Villar Onrubia, D., On the Futures of Technology in Education: Emerging Trends and Policy Implications, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/079734

¹³³ Fengchun, M. & al (2023) Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

¹³⁴ Bourcher, B. (2023) “We cannot fight AI.” How art schools are navigating the challenge of artificial intelligence. Artnet. <https://news.artnet.com/cdn.ampproject.org/c/s/news.artnet.com/art-world/art-schools-and-artificial-intelligence-2377262/amp-page>

¹³⁵ Loos E, Gröpler J, Goudeau M-LS (2023). Using ChatGPT in Education: Human Reflection on ChatGPT's Self-Reflection. *Societies*. 13(8):196. <https://doi.org/10.3390/soc13080196>

¹³⁶ Loos E, Gröpler J, Goudeau M-LS (2023). Using ChatGPT in Education: Human Reflection on ChatGPT's Self-Reflection. *Societies*. 13(8):196. <https://doi.org/10.3390/soc13080196>

¹³⁷ Steijs, J., et al (2023). Comparing the Quality of Human and ChatGPT Feedback on Students' Writing. <https://doi.org/10.33542/osfio/ty3em>

¹³⁸ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

¹³⁹ Tuomi, I., Cachia, R. and Villar Onrubia, D. (2023) On the Futures of Technology in Education: Emerging Trends and Policy Implications, Publications Office of the European Union, Luxembourg, doi:10.2760/079734

¹⁴⁰ Ks. esim. OpenAI: Teaching with AI <https://openai.com/blog/teaching-with-ai>

¹⁴¹ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

¹⁴² Oulun yliopisto (2022) Tekoälyn käyttöä oppimisessa tutkitaan merkittävällä kansainvälisellä rahoituksella. <https://www.oulu.fi/uitiset/tekoalyn-kayttoa-oppimisessa-tutkitaan-merkittavalla-kansainvalisella-rahoituksella-sanna-jarvela>

¹⁴³ University of Oulu (2023) Research to improve human and artificial intelligence interaction. <https://www.oulu.fi/en/news/research-improve-human-and-artificial-intelligence-interaction>

¹⁴⁴ Vrt. Jason M. Lodge, Sujing Yang, Leon Furze & Phillip Dawson (2023) It's not like a calculator, so what is the relationship between learners and generative artificial intelligence?, Learning: Research and Practice, 9:2, 117-124, DOI: 10.1080/23735082.2023.2261106

¹⁴⁵ Fengchun, M. & al (2023) Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

¹⁴⁶ MathGPT (ladattu 19.12.2023) <https://www.mathgpt.com>

¹⁴⁷ European Commission, Directorate-General for Education, Youth, Sport and Culture, Digital education action plan 2021-2027 – Improving the provision of digital skills in education and training, Publications Office of the European Union, 2023, <https://data.europa.eu/doi/10.2766/149764>

¹⁴⁸ Vallionvarainministeriö (ladattu 19.12.2023) EU:n digisaadoksilla luodaan pelisääntöjä digitaalisen ajan toimintaympäristöön <https://vm.fi/eu-n-digisaadokset>

¹⁴⁹ Euroopan komissio (ladattu 19.12.2023) Digitaalisen koulutuksen toimintasuunnitelma 2021-2027. <https://education.ec.europa.eu/fi/focus-topics/digital-education/action-plan>

¹⁵⁰ Tekoälyn ja dataan käyttö opetuksessa ja oppimisessa: <https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en/format-PDF/source-search>

¹⁵¹ Digitaalisen informaatiolukutaidon opas, Faktabaari (2022): www.faktabaari.fi/dil

¹⁵² Vuorikari, R., Kluzer, S. and Punie, Y., DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-48883-5, doi:10.2760/490274

¹⁵³ Ohjeet opettajille ja kasvattajille disinformaation torjumiseksi: <https://op.europa.eu/en/publication-detail/-/publication/a224c235-4843-11ed-92ed-01aa75ed71a1/language-en>

¹⁵⁴ GenAI hanke (ladattu 19.12.2023) Generation AI hanke vastaa kolmeen toisiinsa liittyvään yhteiskunnalliseen haasteeseen. <https://www.generation-ai-stn.fi/>

¹⁵⁵ Esim. Facebookin ryhmä TOT Tekoäly oppimisen tukena https://www.facebook.com/groups/tekoaluyoppimisen_tukena

¹⁵⁶ Digioppimisen areena: <https://www.oph.fi/fi/tapahtumat/2023/digioppimisen-areena-2023-arena-digital-larande-2023>

¹⁵⁷ ITK-konferenssi: <https://itk-konferenssi.fi/>



11. Pedagoginen ajattelu, aktiivinen oppija ja generatiivisen tekoälyn käyttö alakoulussa

TAPANI SAARINEN

Suomessa on painotettu opettajan pedagogista ajattelua, ja tätä osaamista voidaan soveltaa myös generatiivisen tekoälyn, kuten ChatGPT:n käytössä. Alla esitän lyhyesti yhden mahdollisen pedagogisen lähestymistavan tekoälyn käyttöön ja neljä esimerkkiä sen soveltamisesta generatiiviseen tekoölyyn.

Pedagoginen taustateoriani perustuu tekoälyn ja oppimisen yhteensovittamiseen. Se korostaa oppijan roolia aktiivisena, ympäristönsä kanssa vuorovaikutuksessa olevana toimijana. Teoria erottelee oppimisen tilanteisiin, joissa oppija ohjelmoi tietokonetta, ja niihin, joissa tietokone ohjelmoi oppijaa. Pedagoginen taustateoria perustuu muun muassa Seymour Papertin (1928–2016) tutkimuksiin ja pyrkii etsimään mahdolli-

suksia oppimisen yksilöllistämiseen.

Seymour Papert toimi 1960-luvulla kasvatustieteiden professorina ja myöhemmin uralaan tekoälytutkijana MIT:ssä (Massachusetts Institute of Technology). Papert kuvasi opetuksessa käytettävän tekoälyn pedagogiikkaa vuonna 1980 ilmestyneessä klassikkoteoksessaan *Mindstorms*, joka on julkaistu suomeksi nimellä *Lapset, tietokoneet, ajattelemisen taito* (Kirjayhtymä, 1985).

Papertin mukaan tekoäly konkretisoi oppimista, ja siksi voimme hyödyntää omaa oppimistietämystämme ymmärtääksemme tekoälyn toimintaa. Tekoölyyn perehtyminen vie meidät oppimisen peruskysymysten äärelle: Mitä oppiminen itse asiassa on?

Papert korostaa oppimista aktiivisena vuorovaikutuksena ympäristön kanssa, jolloin tietokone voi tarjota oppijalle erilaisia vuorovaikutusympäristöjä. Oppijan aktiivinen rooli korostuu pedagogiikkaa ohjaavassa ajatuksessa, jonka mukaan oppijan tulee ohjelmoida tietokonetta, eikä päinvastoin.

Lasta voi ohjata ymmärtämään tekoälyn toimintaa erilaisten projektien avulla, joita Papert kutsuu nimellä Baby AI. Nämä projektit havainnollistavat jonkin tekoälyyn liittyvän toimintaperiaatteen. Oppija voi esimerkiksi hyödyntää robotissa olevia sensoreita, jotka keräävät tietoa ympäristöstä, ja ohjelmoida robotin liikkumaan luokassa autonomisesti. Näin havainnollistuu jotain itseohjautuvan auton toiminnasta.

Tavoite oppijasta aktiivisena toimijana ja teknologian tuomien mahdollisuuksien hyödyntäjänä voi oman pedagogisen ajattelun mukaan toteutua myös tilanteissa, joissa tekoälyä käytetään ”ohjelmoimaan oppijaa”. Maailmalla yleistyvät esimerkiksi tekoälyä hyödyntävät, adaptoituvat opiskeluympäristöt. Intelligent Tutoring System (ITS) on eräs tekoälyn oppimissovellus, jossa tekoäly hyödyntää oppijasta saatavaa dataa opiskelun yksilöllistämässä.

Pedagogisen ajattelun keskeistä roolia olisi paikallaan konkretisoida myös tekoälyn toiminnassa, kun suomalaisia tekoälyn oppimiskäytön sovelluksia, kuten Intelligent Tutoring System (ITS), kehitetään.

Tekoäly oppimisen apuna

Seuraavaksi esitän neljä esimerkkiä, joissa edellä kuvattu pedagoginen ajattelu konkretisoituu generatiivista tekoälyä, kuten ChatGPT:tä hyödyntävässä oppimistoiminnassa ja oppimisen suunnittelussa. Oppija on esimerkeissä generatiivisen tekoälyn kanssa aktiivisessa vuorovaikutuksessa. Tekoäly tarjoaa ympäristön, jossa tutkiminen ja iteraatio ovat keskeisiä.

Esimerkit koskevat:

- oppimissuunnitelmien tekoa moniälykkyyden teoriaa hyödyntäen
- ajattelun taitojen kehittämistä argumentointia harjoittaen
- tekoälyä tuutorina sekä
- oppijoiden ikäkauteen sopivan tekstitason etsimistä lukemista varten.

Tekoälyä voi pyytää ehdottamaan erilaisia oppimissuunnitelmia. Se voi luoda monia eri vaihtoehtoja, jotka hyödyntävät esimerkiksi moniälykkyyden eri muotoja, kuten Howard Gardnerin Multiple Intelligences -teoriaa. On mielenkiintoista kehittää samasta suunnitelmasta eri versioita, jotka hyödyntävät esimerkiksi kineesteettistä, visuaalista tai musiikillista älykkyyttä.

Moniälykkyyden eri muotojen hyödyntäminen voi palvella opetuksen eriyttämisessä ja oppimisen yksi-

öllistämässä. Tekoälyn avulla generoitu materiaali on yksi esimerkki mahdollisesta lisäarvosta, jota digitaaliset opiskeluympäristöt tuovat oppikirjoihin verrattuna.

Toinen esimerkkini liittyy ajattelun taitojen harjoittamiseen, jossa tekoäly on osoittautunut hyödylliseksi. Olen käyttänyt ChatGPT:tä auttamaan väitteiden ja niiden perusteluiden kehittämisessä.

Tekoäly toimii tehokkaana kumppanina argumenttien tarkastelussa ja luomisessa. Voit testata omia väitteitäsi ja niiden perusteluja ChatGPT:n avulla, pyytää arviota heikoista kohdista ja vahvuuksista.

Tekoälyä voi pyytää luomaan esimerkkejä tietyn tyyppisistä argumenteista, kuten analogiaan perustuvan argumentin, tai asioiden väliseen korrelaatioon perustuvan argumentin (esim. syy-seuraus). Tietyn väitteen perustelua voi näin testata, arvioida ja kehittää eri tavoin esitettynä.

Olen pitkään ollut kiinnostunut ajattelun taitojen harjoittelusta koulussa, erityisesti perustelutaidon kehittämisestä. Tekoäly tarjoaa uudenlaisen työkalun tähän tarkoitukseen.

Opettajana voit hyödyntää tekoälyä ammatillisen kirjallisuuden opiskelussa. Syötä ChatGPT:lle tieteellinen artikkeli tai sen tiivistelmä, ja pyydä sitä erittelemään tekstissä esitetyt pääargumentit ja niiden tukemisen tavat. Saat nopeasti selville, onko argumentaatio vakuuttavaa vai ontuvatko perustelut. Tekoäly auttaa tunnistamaan myös, jos artikkeli perustuu ideologisiin väitteisiin ilman vankkoja perusteluja.

Kolmas esimerkkini käsittelee ChatGPT:n käyttöä tuutorina. Tutkimme alakoulun oppilaiden kanssa esimerkiksi sitä, miten tekoäly voisi auttaa heitä laskulausekkeiden kanssa. Vinkit olivat samansuuntaisia kuin matematiikan kirjassa.

Olen vinkannut tekoälystä myös vanhemmille, mikäli kotona tarvitaan apua oppisisältöön, joka ei enää omilta kouluajoilta muistu mieleen. Matematiikassa tällainen saattaa olla esimerkiksi laskujärjestys.

Mitä enemmän tekoälyä käyttää tuutorina, sitä enemmän voi pyrkiä tuottamaan sen avulla juuri omaan oppimistyyliin ja tavoitteisiin soveltuvia ideoita ja pyytää kehittävästä palautetta.

Tekoäly on tuutorina verraten hyvä palautteen antaja. Oppija voi esimerkiksi syöttää kirjoittamansa kirjoitelman tai esseän ChatGPT:hen ja pyytää arviota sekä kehitysehdotuksia. Myös tämän tekstin kehittämisessä pyydettiin palautetta ja kommentteja sekä kollegoilta että tekoälyltä. Tekoälyä pyysin antamaan ideoita esimerkiksi asioiden tiivistämiseen.

Palautteen antaminen ja saaminen on tärkeä osa oppimisprosessia. Tekoälyn avulla oppija voi saada jopa enemmän henkilökohtaista palautetta kuin mihin opettajalla on mahdollisuus 20–30 oppijan ryhmässä luokassa.

ChatGPT:n julkaisun jälkeen mediassa sai verraten paljon huomiota huoli siitä, että oppilaat generoivat esseensä suoraan tekoälyn avulla. Huoli on aiheellinen

ja muistuttaa aikaa, jolloin Wikipedia oli uusi: Opettajalle saatettiin palauttaa esseitä, jotka oli tehty leikkaa-liimaa-tekniikalla suoraan netistä.

Mutta onko rakentava ratkaisu – nyt tai silloin – kielittää resurssin käyttöä? Vai voisiko koulussa ohjata oppijoita resurssin hyödyllisessä ja omaa oppimista kehittävässä käytössä?

Mahdollisuus hyödyntää sekä inhimillisen tiedon aarreaittaa verkossa että uusia työkaluja, kuten generatiivista tekoälyä, ei nimittäin ole mikään pieni asia.

Neljänneksi, tekoäly pystyy kertomaan perustiedot lähes aiheesta kuin aiheesta, ja antamaan ohjeita erilaisten taitojen harjoitteluun. Riittää kun vain pyytää ja tarkentaa pyyntöjä. Myös itse asioiden oppimiseen tarvitaan oppijan aktiivista toimintaa.

Tekoälyä voi pyytää kertomaan perustiedot eri muodoissa: Olen kokenut hyödylliseksi pyytää tekoälyä kertomaan jonkin asia niin, että 10-vuotias sen ymmärtää (tai muun ikäinen, riippuen siitä minkä ikäisten parissa toimit). Tekoäly kykenee löytämään verraten hyödyllisiä tulokulmia vaikeisiin asioihin: Esimerkiksi ChatGPT:lle tehty pyyntö ”kerro kvanttifysiikasta niin että 10-vuotias sen ymmärtää”, oli hyödyllinen myös lähes 50-vuotiaalle opettajalle tämän asian hahmottamisessa. Vastauksiin on kuitenkin aina syytä suhtautua kriittisesti ja tarkastaa, onko niiden asiasisältö oikea.

ChatGPT:lle voi syöttää valmiin artikkelin ja pyytää sitä muokkaamaan se halutun ikäkauden mukaiselle tasolle.

Aiemmin opettajana on voinut käyttää erilaisia tekstilähteitä ja kertoa sitten suullisesti opittavasta asiasta oppijoiden ikäkautta vastaavalla tavalla ja tasolla. Tekoäly auttaa muokkaamaan valmiin tekstin ikäkauden asettamat vaatimukset huomioon ottaen, jolloin oppijat voivat opiskella lukemalla. Näitä tekstejä voi edelleen hyödyntää opetuksen eriyttämisessä.

Tekoäly on myös verraten hyvä avaamaan käsitteiden tai sanojen merkitystä, mikä palvelee oppijoita, jotka laajentavat sanavarastoaan.

Oppimista voi elävöittää ja yksilöllistää pyytämällä tekoälyä kertomaan opiskeltavan asian esimerkiksi kertomuksen muodossa. Tekoäly keksii hyviä tarinoita vaikkapa veden kiertokulusta. Tekoälyä voi pyytää myös tekemään tietovisan jne. Oma pedagoginen ajattelu ja mielikuvitus on rajana.

Tekoäly perustuu oppimiseen

Edellä esitetyissä neljässä esimerkissä olen havainnollistanut, miten pedagoginen periaate, joka korostaa oppijan aktiivista vuorovaikutusta ympäristönsä kanssa, voi toteutua generatiivisen tekoälyn käytössä. Se toteutuu, kun pedagogisesti ajatteleva opettaja, tai nuori oppija etsii, pohtii, kokeilee ja tutkii erilaisia tapoja hyödyntää tekoälyn tarjoamia mahdollisuuksia tiedon luomiseen ja oppimiseen.

Tekoälyssä on pohjimmiltaan kyse oppimisesta ja oppimiseen liittyvän tiedon tekemisestä konkreettiseksi tekoälyn toiminnassa. Näin ymmärrettynä opettajalla, oppimisen ammattilaisella, on hyvät edellytykset lähestyä ja käyttää tekoälyä. Tekoäly ei ole mitään vierasta oppimiselle, vaan se on itse asiassa oppimisen eräs muoto konkretisoituna teknologiassa.

Opettajan tekoälyä koskeva pedagoginen ajattelu hyötyy omasta kokeilusta, tekoälyn käytön seuraamisesta sekä Suomessa että ulkomailla, kuin myös tutustumisesta tekoälyn käyttöön koulun ulkopuolella. Tämä voi antaa uusia ideoita siitä, miten tekoälyä voisi koulussa hyödyntää.

Tekoälyn koulukäytöstä julkaistaan paljon tutkimustietoa niin artikkeleina kuin kirjoinakin. Tekoälystä käytävän kansallisen keskustelun ohella tekoälyn opetuskäytöstä löytyy monia kansainvälisiä pohdintoja, kuten vaikkapa Unicefin, OECD:n ja Euroopan komission julkaisemat.¹⁵⁸

Omassa ajattelussani tekoälyn opetuskäyttöön liittyy myös kansainvälisen politiikan taso. Koska tekoälyn kehittäminen on keskeinen osa poliittisen hegemonian tavoittelussa, näyttää varmalta, että ainakin tietyt valtiot priorisoivat tekoälyn käytön ja kehittämisen. Koulutuksella pyritään varmistamaan se, että lahjakkaita osajia löytyy kehittämään tekoälyä. Tekoälyn kehittämiseen tarvitaan myös valtava määrä dataa, jonka hankintaan ja käyttöön liittyvät eettiset kysymykset ovat jo nyt tärkeä osa koulutukseen liittyvää keskustelua.

Tekoäly ei ole yksi yhtenäinen asia, vaan joukko erilaisia tekniikoita. Näitä tekniikoita voidaan käyttää opiskelun ja oppimisen apuna monin erilaisin tavoin. Tässä artikkelissa ehdotin, että generatiivisen tekoälyn käyttöä alakoulussa ohjaisi pedagoginen ajattelu, joka voi hyödyntää erilaisia lähtökohtia ja mahdollisuuksia. Tekoäly ei ole vain uusi, hyödyllinen työkalu yhteiskunnassa tärkeän digitaalisen ja teknologisen avaintaidon harjoittelussa, vaan mahdollisuus pohtia, mitä oppiminen perimmältään on, ja kehittää sitä.

¹⁵⁸ Unicefin ohjeistus: <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>;

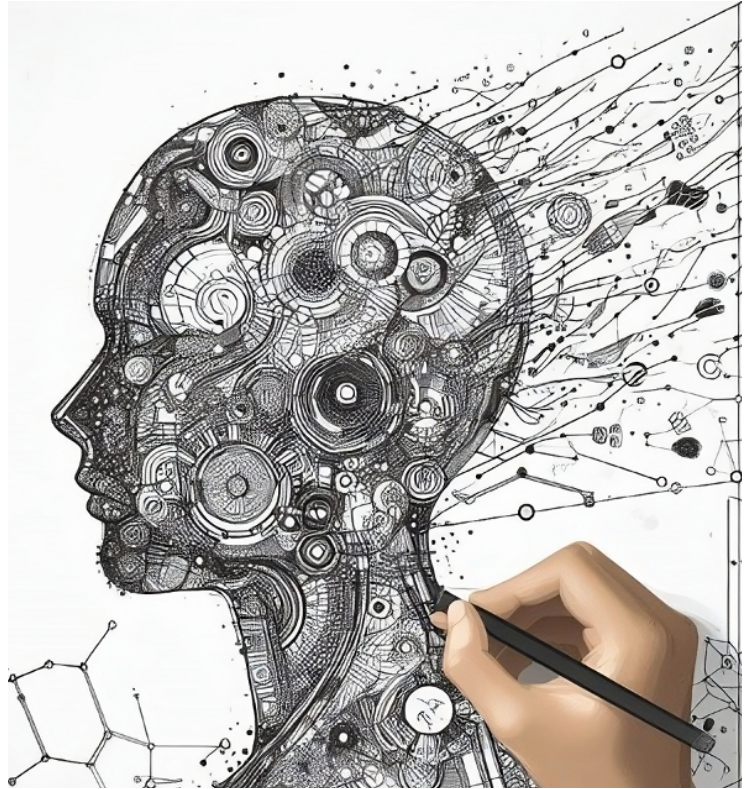
OECD:n pohdintoja: <https://www.oecd.org/education/trustworthy-artificial-intelligence-in-education.pdf>;

Komission laatimia suuntaviivoja: <https://education.ec.europa.eu/news/ethical-guidelines-on-the-use-of-artificial-intelligence-and-data-in-teaching-and-learning-for-educators>

12.

Tekoälysovellukset lukiossa

AKI SAARIAHO



Tekoälysovellukset ilmaantuivat tavallisten kuluttajien käyttöön ChatGPT:n julkistamisen jälkeen vuoden 2022 marras-joulukuun taitteessa. Samalla tavoin kuin kynä ja paperi, painokone, laskin ja tietokone mullistivat vuorollaan tapaamme oppia ja opettaa, olemme nyt saaneet uuden, tehokkaan työkalun, joka ei poistu keskuudestamme, vaikka yrittäisimme rajoittaa sen käyttöä. Parempi strategia on opetella yhdistämään tekoälysovelluksia osaksi opetusta.

Sussexin yliopiston kognitiivisen filosofian professori Andy Clark kuvaa tekoälyn mahdollisuuksia hyvin uudessa kirjassaan *The Experience Machine*, jossa hän lainaa aiemmin Andy Chalmersin kanssa kirjoittamaansa artikkelia "The Extended mind"¹⁵⁹:

Argumentoimme, että yksilön mielen koneiston ei tarvitse rajoittua yksilön aivojen ja keskushermoston koneistoon. Sen ei edes tarvitse rajoittua heidän kehoonsa yleisemmin ymmärrettynä.

Toinen esimerkki on Annie Murphy Paulin kirjasta *The Extended Mind*,¹⁶⁰ jossa hän viittaa samaiseen Andy Clarkiin:

Hän tajusi, että tietokoneesta oli tavallaan tullut osa hänen mieltään, olennainen osa hänen ajattelu-prosessiaan. Kannettavan tietokoneen käyttö laajensi tehokkaasti hänen henkistä kapasiteettiaan, jolloin hänen aivonsa pystyivät ylisuorituksiin – ajattelemaan tehokkaammin ja tuloksellisemmin, älykkäämmin, kuin mihin ne pystyivät ilman laitetta. (...) Hänen aivonsa plus tietokoneensa olivat yhtä kuin hänen mielensä, laajennettuna.

Omassa opetustyössäni toisella asteella ja etenkin lukiossa tekoälysovellusten käyttöä rajaavat useat reu-

naehdot. Käyn tässä läpi useita niistä, mutta tuskin kaikkia. Lukiossa suurin kysymys liittyy opiskelun tarkoitukseen: opiskellaanko siellä elämää vai ylioppilaskirjoituksia varten? Koska ei ole tarkoituksenmukaista valita näiden välillä, tulisi tekoälyn käytön ja valittujen sovellusten avulla kurseilla kumpaakin tavoitetta. Opiskelijoita kannattaa valmentaa sekä yo-kirjoituksia että jatko-opintoja ja työelämää varten.

Yo-kirjoitusten vuoksi tekoälysovelluksia tulisi käyttää mahdollisimman hyvään pedagogiseen tarkoitukseen. Ei riitä, että saadaan tekoälyavusteisesti hyvää käännöstä tai kuvanmuokkausta. Opiskelijoilla tulisi teettää monipuolisia töitä, jotka haastavat heitä ja palvelevia pitkään. Kurssitöissä pitäisi mielestäni olla mahdollisimman paljon transversaaliutta, opittujen taitojen siirtämistä oppiaineesta toiseen.

Opiskelijoiden aika on jaettu niin moneen paikkaan, että opetussuunnitelmissa toivotut poikkitieteelliset ja systeemiset kurssit, joissa yhdistettäisiin oppiaineita ja mahdollisesti opettajia, eivät toteudu koska ne eivät mahdu lukujärjestyksiin. Näitä olisi kuitenkin mahdollista toteuttaa katsomalla koulun kurssitarjontaa kokonaisuutena ja tarjoamalla erilaisia palasia erityyppisistä taidoista eri puolilla opiskelijoiden oppipolkua. Tässä tekoälysovellukset olisivat oiva apu.

Esimerkiksi muuttuvaa elinympäristöä voisi tutkia hyödyntämällä dataa kaupunkisuunnittelusta, Tilastokeskuksesta ja avoimen datan lähteistä. Opiskelijat voisivat käsitellä dataa tekoälysovelluksilla luettavaan ja raportoitavaan muotoon matematiikassa ja fysiikassa. Saman datan tuloksia voisi tarkastella elinympäristön muutoksina biologissa ja maantieteessä. Elinympäris-

tön muuttaminen ihmisen toiminnan seurauksena tarjoaa näkökulmia filosofiaan ja psykologiaan. Kun tuloksia yhdistelee, opiskelijat voisivat kirjoittaa niiden pohjalta suomen kielessä tai median ja vaikuttamisen opinnoissa mielipidekirjoituksen tiedotusvälineeseen. Tekoälysovelluksia voi hyödyntää työskentelyssä kaiken aikaa tekstien jalostamiseen sekä tilastojen prosessointiin.

Mitä Napoleon ajatteli Waterloo jälkeen?

Tekoälyn käytössä pidän opiskelijoiden tärkeimpinä taitoina kirjoitettujen ja puhuttujen syötteiden muodostamista sekä oman opiskelumotivaation harjaannuttamista. Opiskelijoiden täytyy itse päätyä siihen tulokseen, että tärkein taito on oppia oppimista. Kun he oppivat koulumaan tekoälyä ja työskentelemään sen kanssa, on opiskelu mielekästä jatkossakin.

Syötteiden tuottamisen tulisi olla kansalaistaito. Tekoälysovellusten järjestyvä käyttö taas ratkaisee, käyttääkö opiskelija aikansa oppimiseen ja kehittymiseen vai helppoon tiehen, kurssien läpäsyyntä. Tämä on aina ollut opettajan työn suola ja todellinen koetin – minusta se vain konkretisoituu nyt entisestään. Kun työkalut tekevät nykyisten tehtävien tekemisestä vastenmielisiä ja toisaalta houkuttelevan helppoa napin painalluksella, opettajan tulee muotoilla tehtävät toisin.

Opiskelijoiden on ennen kaikkea ymmärrettävä, mihin tekoälysovelluksia kannattaa käyttää ja mihin ei. Esimerkiksi generatiiviset kielimallit eivät ole hakukoneita, vaikka houkutus käyttää niitä siihen tarkoitukseen on suuri. Lähteiden tarkistaminen ja faktantarkistus ovat taas äärettömän tärkeitä taitoja, joita pitäisi painottaa koko ajan.

Seuraavaksi annan muutaman esimerkin aineenopeutuksen kontekstista. Opettajan näkökulmasta tekoälysovellukset vaikuttavat jonkin verran sopiviin tehtävyytyyppeihin ja kurssitöiden relevanttiteuteen opiskelijoille. Toki vastaavia tehtäviä olisi voinut teettää jo aiemmin, mutta tekoälyn avulla se on helpompaa ja mielikuvituksellisempaa.

Ensimmäinen tehtävä käsittelee satunnaisesti simulaatiota historian hahmosta. Pyysin ChatGPT:tä esittämään Napoleonin, jota ”haastattelin” Waterloo taistelun jälkeen. Tein tämän naiivin esimerkin, koska Napoleonin hahmo on tunnettu ja Waterloo taistelusta useimmilla meistä jo ennestään käsitys. Kielenä käytin englantia, koska se on ChatGPT:lle helpompaa. Sama onnistuisi todennäköisesti suomeksi hyvin, mutta harvinaisempien esimerkkien tai tilanteiden kohdalla ei välttämättä yhtä hyvin.

Erityisesti kielenopetuksessa tähän liittyy monia asioita, joita opetella opiskelijoiden kanssa. Miten opiskelijoiden tekoälylle antamat pyynnöt toimivat eri kielillä? Kuinka tarkasti opiskelijat osaavat muotoilla syötteitä eri kielillä?

You

tehdään simulaatio. Ole sinä Napoleon ja minä olen London Times -lehden toimittaja. Olet juuri hävinnyt Waterloo taistelun. Miltä tuntuu tämän taistelun jälkeen ja mitä aiot seuraavaksi? Ole aika tiivis vastauksessasi ja sisällytä mukaan Napoleonin suunnitelmien laajempia kaaria.

ChatGPT

Hyvä toimittaja,

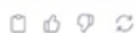
Waterloon tappio on kieltämättä kova isku. Se on merkki siitä, että minun täytyy miettiä uudelleen paikkaani Euroopan ja maailman näyttämöllä. Tämän hetken tuntemukseni ovat sekaisin – pettymystä, yllätystä, mutta myös pohdintaa siitä, miten strategiani ja tavoitteeni ovat muuttuneet.

Seuraavaksi on edessä mahdollisesti maanpako. Tämä voi olla tilaisuus pohtia ja kirjoittaa muistelmiani, jakaa opetuksiani ja näkemyksiäni tuleville sukupolville. Vaikka minun on ehkä väistyttävä poliittisesta elämästä, aion jatkaa vaikuttamista ajatuksillani ja kirjoituksillani. Historian virta ei pysähdy, ja vaikka en ole enää sen keskiössä, minulla on edelleen rooli pelattavana.

Lopulta, vaikka Waterloo tappio merkitsee loppua eräälle luvulle historiassani, se ei ole loppu minun tarinalleni. Aion käyttää tämän kokemuksen oppimiseen ja kasvuun.

Terveisin,

Napoleon Bonaparte



Napoleonin käsittelyt roolipelisimulaatio oli sinänsä hassuttelujuttu. "Napoleonin" vastaukset kuitenkin avasivat ikkunan tärkeämpiin aiheisiin. Niissä näkyi vihjeitä Napoleonin haaveista vaikuttaa edelleen maailmanpolitiikkaan ja hänen aikomuksistaan kirjoittaa muistelmia. "Napoleonin" vastauksia voisi käyttää oppimisen syventämiseen. Esimerkkiä voisi myös jatkaa Ranskan historian muilla hahmoilla tai Ranskan valtion kehityksellä Napoleonin jälkeen, aina Euroopan unionin muodostumiseen asti. Miten Napoleon näki Euroopan yhtentymiskehityksen?

Tämä on tekoälyavusteisen opetuksen tätä päivää. Oletan, että piakkoin voimme VR-lasien avulla tavalta samaisen "Napoleonin" ja käydä tuon keskustelun ennen virtuaalista Waterloo taistelua tai sen jälkeen - tai simuloita "Napoleonin" hallitseman Ranskan parlamentin istuntoa VR- tai Metaverse-todellisuudessa avatar-hahmoilla. Tämä avaisi eri oppiaineiden sisältöihin syvempää kokeilua tekoälysovellusten avulla ja mahdollistaisi opiskelijoiden kontakteja kansainvälisesti, ympäristön kannalta kestävämmiin ja matka-aikaa säästäen.

Filosofian tehtäviä tekoälyllä

Toinen esimerkkini käsittelee koekysymysten laatimista. Lokakuussa 2023 syötiin ChatGPT:lle filosofian etiikan opintojakson 2 LOPS2019 sisällön. Pyysin ChatGPT:tä antamaan kolme koekysymystä ja kolme kurssityötä.

Tarkennuksena, että ChatGPT:lle voi antaa täsmällisiä ohjeita siitä, miten sen toivotaan vastaavan ja minkä tyyppisiä sisältöjä vastaukset käsittelevät. Olen kirjoittanut omiin kuvauksiini ilmastonmuutoksen ja filosofian kiinnostuksen kohteikseni sekä osaamisekseni, mikä näkyy sen vastauksissa. ChatGPT vastaa tämän vuoksi aina niin, että olettaa minun haluavan filosofian ja ilmastonmuutoksen ratkaisemiseen liittyvien asioiden näkyvän vastauksissa.

Vastaukset olivat aivan riittävän hyviä muutamaan kysymykseen ja kurssityön aiheiksi. Noista voisi hyvin jatkaa ja kehittää itse eteenpäin tai syventää kysymällä ChatGPT:ltä lisää, tai pyytää tekoälyltä apua pisteytysohjeiden laatimisessa.

Kannattaa huomata, että ChatGPT:n käyttöä on toistaiseksi rajoitettu monessa kunnassa ja kaupungissa tietoturvasyistä ja EU:n tietosuojasetuksen GDPR:n vuoksi. Esimerkiksi Bing on sen sijaan vapaammin käytössä.

Käsittele moraalista objektivismia, relativismia ja subjektivismia. Miten nämä näkemykset eroavat toisistaan, ja mitä eettisiä kysymyksiä ne herättävät? Anna esimerkki kustakin näkemyksestä.

2. Koekysymys 2:

Tarkastele moraalin luonnetta normijärjestelmänä ja sen eroa lakiin ja tapoihin. Mikä rooli on moraalilla yksilön päätöksenteossa, ja miten se liittyy yhteiskunnan eettisiin kysymyksiin?

3. Koekysymys 3:

Keskustele eettisistä kysymyksistä liittyen ympäristöön, erityisesti ilmastonmuutokseen ja eläinten oikeuksiin. Miten filosofinen etiikka voi auttaa meitä ymmärtämään ja ratkaisemaan näitä globaaleja haasteita?

Kurssityöaiheet:

1. Kurssityöaihe 1:

Tutki hyve-etiikkaa ja sen soveltamista ihmissuhteisiin. Valitse yksi hyve, kuten rehellisyys tai myötätunto, ja tarkastele, miten se vaikuttaa ihmisten välisiin moraalisiin valintoihin. Esitä konkreettisia esimerkkejä.

2. Kurssityöaihe 2:

Analysoi filosofisia teorioita elämän hyvyydestä ja merkityksellisyydestä. Vertaa esimerkiksi hedonismia ja eudaimonismia ja arvioi niiden soveltuvuutta moderniin elämäntapaan.

3. Kurssityöaihe 3:

Tutki ympäristöä ja luontoa koskevia eettisiä kysymyksiä syvemmin. Fokusoi ilmastonmuutokseen ja sen moraalisiin ulottuvuuksiin. Ehdota konkreettisia eettisiä toimenpiteitä, jotka voivat auttaa lieventämään ilmastonmuutoksen vaikutuksia.

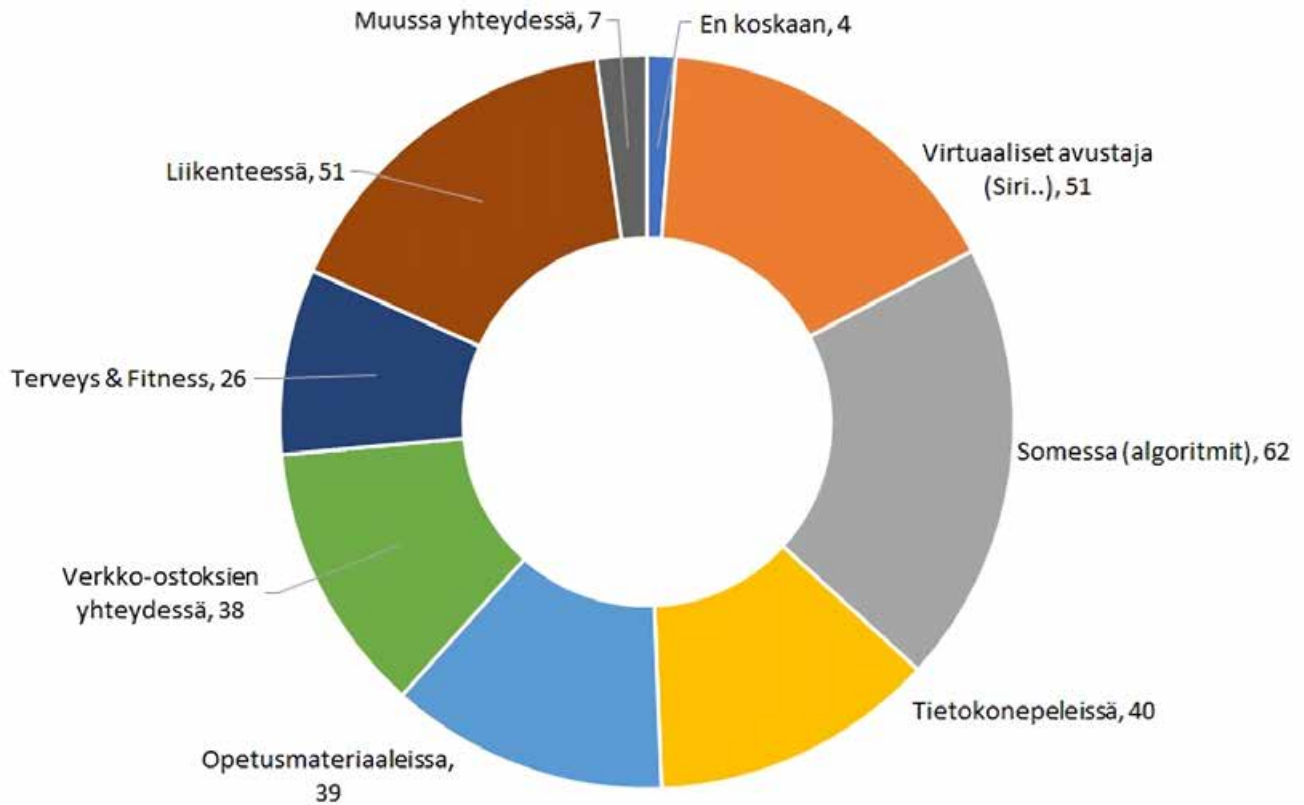
 Regenerate

Kysely lukiolaisille

Lopuksi esittelen opiskelijoiden omia ajatuksia tekoälyn käytöstä. Lokakuussa 2023 tein 145 lukiolaiselle kyselyn, johon sain 75 vastausta:

Oletko käyttänyt AI:ta jokapäiväisessä arjessa. Jos kyllä, niin missä yhteydessä?

N=75



Kun kysyin opiskelijoiden mielipidettä tekoälyn tuomiin eettisiin haasteisiin, opiskelijoita askarruttivat esimerkiksi seuraavat asiat:

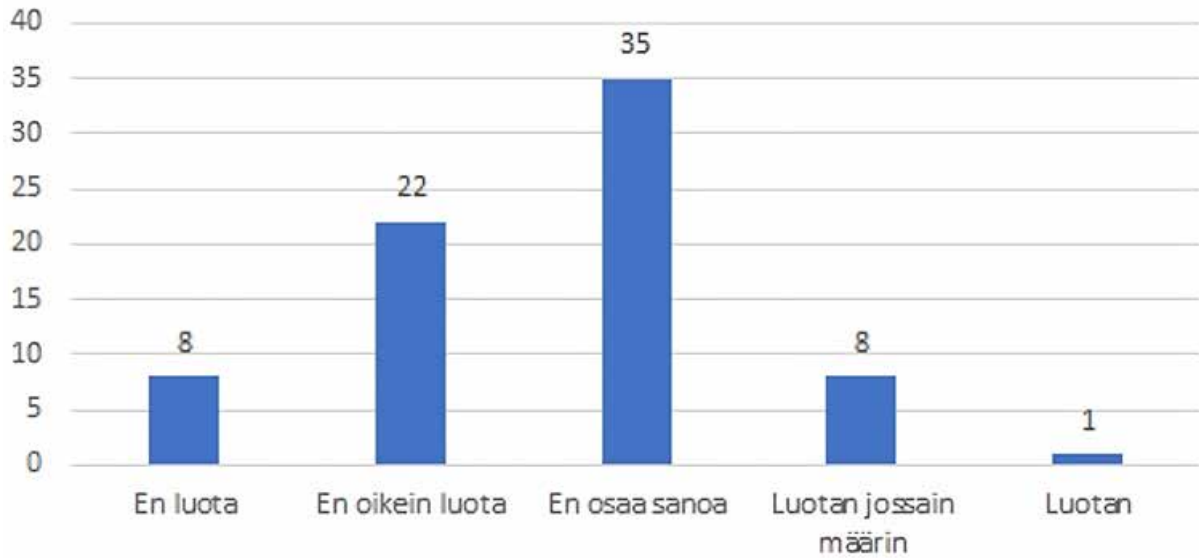
- Mieluusti ei varastettaisi taiteilijoiden tekemiä kuvia tai musiikkia kuva- ja musiikkigeneraattoreita varten.
- Lähdemateriaalin eettisyys (kysytään lupa).
- Haluaisin, että AI pystyisi tekemään tieteellisiä tutkimuksia ja tuotekehittelyä, koska se tietää enemmän ja pystyy ajattelemaan kaikkia mah-

dollisia osa-alueita nopeammin ja tarkemmin kuin yksikään ihminen.

- Sodat voitaisiin siirtää tekoälyn kehittämille pelialustoille.
- AI voi korvata ihmiset opettajina, mutta en usko sen olevan kannattavaa, koska oikean ihmisen opetus innostaa enemmän opiskelemaan.
- Tekoäly saisi keksiä myös hoidon syöväille.

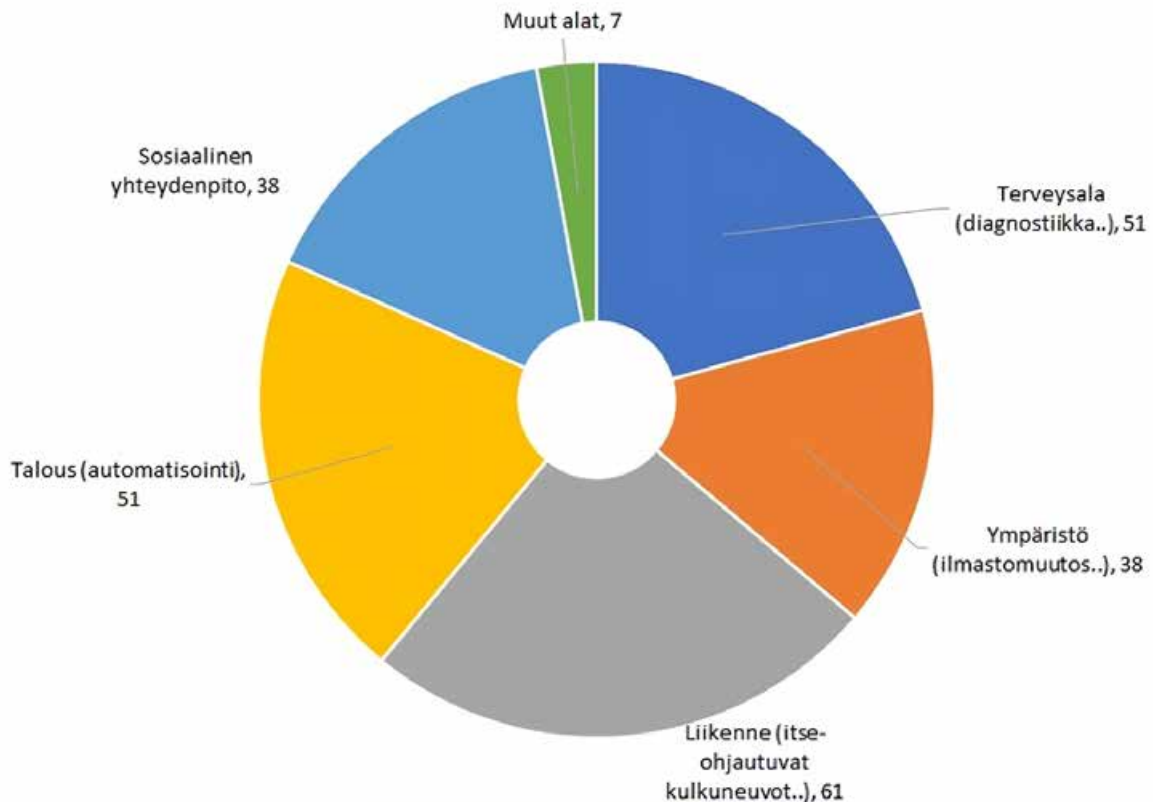
Kysyin myös, luottavatko opiskelijat tekoälyyn:

Kuinka paljon luotat siihen, että tekoäly teksisi elämäsi vaikuttavia päätöksiä N=74



Lopuksi kysyin vielä lukio-opiskelijoiden odotuksia tekoälyn suhteen seuraavien kymmenen vuoden osalta:

Mille aloille mielestäsi tekoälyllä tulee olemaan merkittävä vaikutus? Voit valita usean vaihtoehdon. N=75



Kaiken kaikkiaan nuorilla tuntuu olevan samansuuntaisia ajatuksia ja odotuksia tekoälyavusteisen tulevaisuuden suhteen kuin aikuisillakin. Osa tietää hyvin jo

nyt missä tekoälyä on ja miten se toimii, osa ei vielä hahmota, missä kaikkialla tekoälyä on jo integroitu elämäämme.

¹⁵⁹ JSTOR-artikkelitarkistus: <http://www.jstor.org/stable/3328150>, käyty 20.12.2023.

¹⁶⁰ Annie Murphy Paul: The Extended Mind – The Power of Thinking Outside the Brain. Houghton Mifflin Harcourt, 2021.

13. Pysäyttääkö tekoäly osaamisemme kehityksen?



MATT CRONIN

Kuvittele, että sinulla on taikalaatikko. Jos kosket laatikkoa, se luo sinulle välittömästi tuotoksen, joka on laadultaan parempi kuin kaikkien muiden tuotokset. Koskisitko sitä? Kuinka usein? Kuvittele nyt, että kaikilla on tämä taikalaatikko. Mitä kannustimia se loisi? Miten nämä kannustimet muokkaisivat ihmiskunnan tulevaisuutta?

Nämä ovat kriittisiä kysymyksiä, joihin ChatGPT ja muut generatiiviset tekoälysovellukset (AI) vaativat meitä vastaamaan lähiaikoina.¹⁶¹ Tekoälysovellukset käyttävät neuroverkkoja tuottamaan räätälöityjä tuotteita syötteiden perusteella. Tällä hetkellä generatiivinen tekoäly luo vakuuttavia yliopistollisia esseitä, tietokonekoodia ja taidetta. Se on jo voittanut kansainvälisen valokuvakilpailun, tehnyt hittisinglen ja saanut huippupisteitä korkeakoulujen pääsykokeissa.¹⁶²

Ja se kehittyy nopeasti.¹⁶³ Generatiivisen tekoälyn käyttöä on myös yhä vaikeampi havaita¹⁶⁴, sillä sen dramaattinen kehitys ja älykkäät käyttäjäkehotukset poistavat vivahteita, joiden avulla ihmiset ja ohjelmat voivat tunnistaa sen käytön.¹⁶⁵

Vaikka generatiivisen tekoälyn laajamittainen käyttöönotto tuo lupaavia etuja, niin monet ovat oikeutetusti huolissaan sen riskeistä. Asiantuntijat ovat arvioineet, että tekoäly korvaa noin 25 prosenttia työpaikoista Yhdysvalloissa ja EU:ssa – tosin samalla syntyy jonkin verran myös uusia työpaikkoja.¹⁶⁶

Vaikka tekoäly on kehitetty ajamaan etuamme, niin lajimme kannalta on kuitenkin olemassa todellinen riski: mitä ihmiskunnalle tapahtuu, kun toisiaan seuraavat sukupolvet oppivat vain koskemaan tekoälyn ”taikalaatikkoa” ja ovat täysin taitamattomia ilman sen apua?

Minkä tahansa taidon oppiminen tai aiheen syvälinen ymmärtäminen vaatii paljon aikaa¹⁶⁷ ja vaivaa¹⁶⁸ lahjakkaaltakin oppijalta. Vaikka oppimisen prosessi on välillä stressaava, niin se on loppujen lopuksi palkitseva, vaikka se sisältää epäonnistumisen riskin ja vie usein tuhansia tunteja¹⁶⁹. Ensimmäistä kertaa ihmiskunnan historiassa kokonainen sukupolvi voi ohittaa tämän prosessin ja silti edetä (ainakin jonkin aikaa) koulussa

ja työssä. He voivat koskettaa taikalaatikkoa ja saada aikaan loistavia tuloksia ilman vaivannäköä. Tämä on tavattoman houkutteleva mahdollisuus.

Kuten useimmilla faustisilla sopimuksilla, myös generatiivisen tekoälyn hyödyntämisellä on kuitenkin piilevä hinta. Aina kun kosketat laatikkoa, et todellakaan opi – ainakaan tavalla, josta olisi sinulle merkittävää hyötyä. Kehität tekoälyn neuroverkkoa, et omaasi.

On olemassa pelko siitä, että ajan myötä epäpätevyytemme lisääntyy tekoälyn turvautuessa. Tämä kehityskulku johtaa tilanteeseen, jossa emme pysty elämään ilman tekoälyn tuomaa tukea. Riippuvuus tekoälystä – erityisesti henkilökohtaisen ja ammatillisen kehityksen kriittisinä aikoina – riistää yksilöiltä henkisen elinvoiman ja mahdollisuuden todelliseen oppimiseen. Meistä tulee oman luomuksemme orjia, emmekä kykene ajattelemaan tai toimimaan järkevästi ilman sen apua.

Digitaalinen vesileima käyttöön

Tekoälyn kannattajat ovat sitä mieltä, että tekoälyn käyttö vain muuntaa oppimisen luonnetta, eikä suinkaan vähennä ihmisen oppimista ja osaamista. Tämä näkemys on ensikuulemalta houkutteleva. Me olemme kehittäneet kognitiivisia apuvälineitä, kuten lukutaidon, kirjapainon, internetin, jotka ovat mullistaneet yhteiskuntaa mutta hyödyttäneet lopulta käyttäjää. Tekoäly on erilainen. Toisin kuin kirjat ja verkkosivustot, tekoäly ei pelkästään auta löytämään ja oppimaan tietoa. Se tekee työn puolestasi.

Kirjan lukeminen ei johda siihen, että pöydällesi ilmestyy essee. Kameran käyttäminen ei luo välittömästi mitä tahansa kuvaa. Generatiivinen tekoäly tekee juuri niin. Se on verrattavissa siihen, että menet kuntosalille ja annat jonkun toisen nostella painoja puolestasi ja odotat, että omat lihaksesi vahvistuvat. Tämä on absurdia ja samalla hyvä argumentti tekoälyn käytön rajoittamiselle tietyissä oppimistilanteissa.

Yksinomaan tekoölyyn luottaminen opiskelussa, työssä tai kriittisessä ajattelussa tekee meistä haavoittuvia. Se, joka hallitsee tekoölyalustoja ja (myös itse tekoölyä), hallitsee tehokkaasti ihmiskuntaa.

Luottamus tekoölyyn lisää myös kyberriskejä. Mitä tapahtuu, kun tekoölyn avustuksella leikkauksia suorittavalla lääkärillä ei ole internet-yhteyttä tai tekoölyn käytössä ilmenee teknisiä häiriöitä?

Meidän on myös pohdittava mielenterveysriskejä ennen kuin sitoudumme tekoölyyn. Nuoremmat sukupolvet ovat jo alttiita huijarisyndroomalle¹⁷⁰, jonka seurauksena masennus ja itsemurhat ovat lisääntyneet.

Tekoölyn käytöllä on myös yhteys luovuuteen ja innovaatioon. Generatiivinen tekoöly tuottaa sisältöä vain ihmisen tekemän työn pohjalta.¹⁷¹ Voiko tekoöly päätellä fysiikan seuraavan perustavanlaatuisen muutoksen tämän tietämyksen perusteella vai tarvitaanko siihen nerokkaan ihmisielen intuitiivinen hyppy? Vaikka emme tiedä vastausta tähän kysymykseen, meidän pitäisi varmistaa, että tämän vuosisadan Einsteinilla on kyky istua ja pohtia syvällisesti perustavanlaatuisia toituksia.

On selvää, että nämä riskit eivät tarkoita sitä, että tekoölystä pitäisi luopua. Se on sekä väärin että – Molokin ongelman¹⁷² vuoksi – käytännössä mahdotonta. Sen sijaan meidän on varmistettava, että tekoölyn kehittäminen ei estä ihmisen älykkyyden kehitystä pitkällä aikavälillä.

Tekoölyasiantuntijat ovat oikeutetusti huolissaan tekoölyn asianmukaisesta käytöstä, sillä huonosti ymmärretyt toimet voivat johtaa siihen, että tekoöly vahingoittaa ihmiskuntaa. Vaikka tekoölyn käytön sopeuttaminen onnistuisikin, ihmisten kannustinjärjestelmien kohdentamisen laiminlyönti voi johtaa synkkään ja tietämättömään tulevaisuuteen.

Me emme ottaneet tätä ratkaisevaa askelta sosiaalisen median algoritmien kanssa, ja seuraukset ovat nyt nähtävissä.¹⁷³ Se on vain pieni osoitus niistä riskeistä, joita tällä hetkellä kohtaamme. Siksi meidän on varmistettava, että kannustamme ihmisiä mukautumaan

tekoölyyn asianmukaisesti, jotta sitä käytettäisiin lajia hyödyntävällä tavalla.

Ihmisen kriittisen ajattelun turvaaminen edellyttää useita toimia. Tässä on ensimmäinen ehdotus: Vaaditaan, että kaikki tekoölyalustat ottavat käyttöön digitaaliset vesileimat¹⁷⁴, jotka takaavat tekoölyn tuotosten tunnistettavuuden.¹⁷⁵

Vaikka OpenAI on julkistanut nokkelan vesileimajärjestelmän¹⁷⁶, niin sen vapaaehtoinen käyttöönotto ei riitä. Ilman valtiollisia vaatimuksia markkinat kannustavat tarjoamaan vesileimattomia palveluja. Digitaalisten vesileimojen yleinen käyttöönotto antaisi kouluille ja työnantajille mahdollisuuden tunnistaa tekoölyn avulla tehdyt tuotokset. Tämä puolestaan kannustaisi käyttämään tekoölyä lähtökohtana tai ”avustajana” sen sijaan, että se korvaisi kriittisen ajattelun.

Digitaaliset vesileimat auttaisivat myös ihmiskuntaa sen kilpajuoksussa tekoölyn levittämää disinformaatiota vastaan, ja ne voisivat tehdä tunnistettavaksi tekoölyn läpinäkymättömän toiminnan. Kyseessä on yksinkertainen ja järkevä toimenpide, jonka jopa kaikkein kiihkeimpien tekoölyn puolestapuhujien tulisi hyväksyä.

Olemme nimenneet itsemme lajina Homo sapiensiksi, ”viisaaksi ihmiseksi”. Nyt on tullut hetki, jolloin meidän on todistettava, että olemme tämän nimen arvoisia. Jos emme varmistaa tekoölyn taikalaahtikon vastuullista käyttöä, emme ainoastaan osoita omaa tietämättömyyttämme, vaan tuomitsemme lapsemme kasvamaan tekoölyn hallitsemaan epäpätevien ja tietämättömien ihmisten yhteiskuntaan. Meidän on pohdittava kriittisesti tätä asiaa ja toimittava päättäväisesti nyt, ennen kuin riippuvuus omasta keksinnöstämme vie meiltä tämän kyvyn lopullisesti.

Käännös ja muokaus alkuperäisestä artikkelista kirjoittajan luvalla:

<https://thehill-com.cdn.ampproject.org/c/s/thehill.com/opinion/technology/4007320-do-advances-in-ai-risk-a-future-of-human-incompetence/amp/>

¹⁶¹ OpenAI (ladattu 19.12.2023) Introducing ChatGPT <https://openai.com/blog/chatgpt>; https://en.wikipedia.org/wiki/Generative_artificial_intelligence.

¹⁶² Wikipedia (ladattu 19.12.2023) https://en.wikipedia.org/wiki/Generative_artificial_intelligence. Essee: <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>; tietokonekoodi: <https://blog.tryamigo.com/gpt-4-coding/>; taide: <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>; hitsisingle: <https://time.com/6273529/drake-the-weeknd-ai-song/>; valokuvakilpailu: <https://www.bbc.com/news/entertainment-arts-65296763>; pääsykokeet: <https://arstechnica.com/information-technology/2023/03/openai-announces-gpt-4-its-next-generation-ai-language-model/>

¹⁶³ OpenAI (2023) GPT-4 Technical report <https://cdn.openai.com/papers/gpt-4.pdf>

¹⁶⁴ Zhou, C. et al. (2022) Paraphrase identification with deep learning: A review of datasets and methods. Arxiv. Cornell University. <https://arxiv.org/abs/2212.06933>

¹⁶⁵ Nolasco, N. (2023) Simple method to beat zeroGPT with no edits to ChatGPT generated output. Medium <https://blog.openai.com/simple-method-to-beat-zeroGPT-with-no-edits-to-chatgpt-generated-output-485bbaba8707>

¹⁶⁶ Hatzius, J. (2023) The potentially large effects of artificial intelligence on economic growth. Goldman Sachs economic research. https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf

¹⁶⁷ Wong NC. (2015) The 10 000-hour rule. Can Urol Assoc J. 9(9-10):299. doi: 10.5489/cuaj.3267

¹⁶⁸ Cherry, K. (2022) How hard is it to become an expert? Verywell mind <https://www.verywellmind.com/expertise-how-hard-is-it-to-become-an-expert-at-something-4173614>

¹⁶⁹ Ericsson, A. & Pool, R. (2017) Peak – how all of us can achieve extraordinary things. <https://parentotheca.com/2021/06/14/peak-anders-ericsson-book-summary/>

¹⁷⁰ Eruteya, K. (2022) You're not an imposter. You're actually pretty amazing. HBR. <https://hbr.org/2022/01/youre-not-an-imposter-youre-actually-pretty-amazing> ; Wikipedia (loaded 19.12.2023) https://en.wikipedia.org/wiki/Impostor_syndrome

¹⁷¹ Chiang, T. (2023) ChatGPT is a blurry JPEG of the web. The New Yorker. <https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web>

¹⁷² Miessler, D. (ladattu 19.12.2023) Moloch: the most dangerous idea <https://danielmiessler.com/p/moloch-the-most-dangerous-idea/>

¹⁷³ United Nations (2023) "Urgent need" for more accountability from social media giants to curb the hate speech. <https://news.un.org/en/story/2023/01/1132232>

¹⁷⁴ Wikipedia (ladattu 19.12.2023) https://en.wikipedia.org/wiki/Digital_watermarking

¹⁷⁵ Heikkilä, M. (2023) A watermark for chatbots can expose text written by an AI. MIT technology review. <https://www.technologyreview.com/2023/01/27/1067338/a-watermark-for-chatbots-can-spot-text-written-by-an-ai/>

¹⁷⁶ Collins, K. (2023) How can ChatGPT embed watermark in the text it generates. The New York Times. <https://www.nytimes.com/interactive/2023/02/17/business/ai-text-detection.html>

LIITE 1

Unescon ohjeistus koulutuksen järjestäjille

Unesco on julkaissut konkreettisia suosituksia poliittisille päättäjille ja koulutuksen järjestäjille tekoälypohjaisten ohjelmistojen käytön suunnittelusta ihmisten toimijuuden edistämiseksi ja opiskelijoiden, opettajien ja tutkijoiden tukemiseksi.

Ohjeissa ehdotetaan seitsemää keskeistä toimenpidettä tekoälyn käyttöön koulutuksessa:

Vaihe 1: Hyväksytään kansainväliset tai alueelliset yleiset tietosuoja-asetukset tai laaditaan kansalliset tietosuoja-asetukset. GenAI-mallien kouluttaminen on edellyttänyt kansalaisten verkkotietojen keräämistä ja käsittelyä monissa maissa. Tietojen ja sisällön käyttö ilman suostumusta haastaa tietosuojaa entisestään.

Vaihe 2: Hyväksytään/uudistetaan ja rahoitetaan tekoälyä koskevia kansallisia strategioita. Generatiivisen tekoälyn sääntelyn on oltava osa laajempia kansallisia tekoälystrategioita, joilla voidaan varmistaa tekoälyn turvallinen ja tasapuolinen käyttö kaikilla salloilla, myös koulutuksessa.

Vaihe 3: Vakiinnutetaan ja pannaan täytäntöön tekoälyn etiikkaa koskevat erityissäännökset. Jotta tekoälyn käytön aiheuttamat eettiset näkökohdat voidaan ottaa huomioon, tarvitaan erityisiä säännöksiä.

Vaihe 4: Mukautetaan nykyisiä tekijänoikeuslakeja tai pannaan ne täytäntöön tekoälyn tuottaman sisällön sääntelemiseksi: Tekoälyn yleistymisen on tuonut uusia tekijänoikeudellisia haasteita, jotka koskevat sekä tekijänoikeudellisesti suojattua sisältöä tai teoksia, joilla malleja koulutetaan, että niiden tuottaman ”ei-inhimillisen” tiedon asemaa.

Vaihe 5: Generatiivista tekoälyä koskevien sääntelypuitteiden laatiminen: Tekoälyteknologioiden nopea kehitysvauhti pakottaa kansalliset ja paikalliset hallintoviranomaiset nopeuttamaan sääntelyn uudistamista.

Vaihe 6: Rakennetaan valmiuksia generatiivisen tekoälyn asianmukaiseen käyttöön koulutuksessa ja tutkimuksessa: Koulujen ja muiden oppilaitosten on kehitettävä valmiuksia ymmärtää GenAI-työkalujen mahdolliset hyödyt ja riskit.

Vaihe 7: Pohditaan GenAI:n pitkän aikavälin vaikutuksia koulutukseen ja tutkimukseen: GenAI:n vaikutus ja seuraukset tiedon luomiseen, siirtämiseen ja validointiin – opetukseen ja oppimiseen, opetussuunnitelmien suunnitteluun ja arviointiin sekä tutkimukseen ja tekijänoikeuksiin.

Ohjeistus perustuu koulutuksessa humanistiseen lähestymistapaan, jolla edistetään ihmisen toimijuutta, osallisuutta, tasa-arvoa, sukupuolten tasa-arvoa, kulttuurista ja kielellistä monimuotoisuutta sekä moniarvoisia mielipiteitä ja ilmaisua. Ohjeet ovat yhdenmukaisia Unescon tekoälyn etiikkaa koskevan suosituksen¹⁷⁷ sekä Pekingin konsensus tekoälystä ja koulutuksesta -asiakirjan¹⁷⁸ kanssa, ja ne vastaavat myös ”Koulutuksen uutta yhteiskuntasopimusta”¹⁷⁹, jossa kehoitetaan määrittelemään uudelleen ihmisten ja teknologian välinen suhde.

¹⁷⁷ UNESCO (2022) Recommendation on the ethics of artificial intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

¹⁷⁸ UNESCO (2019) Beijing consensus on artificial intelligence in education. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>

¹⁷⁹ UNESCO (2021) Reimagining our futures together: a new social contract for education. <https://unesdoc.unesco.org/ark:/48223/pf0000379707.locale=en>

LIITE 2

Tekoälyn käytön keskeiset eettiset kysymykset opettajille

Koulujen johdon ja opettajien on hyvä pystyä asettamaan tarkentavia kysymyksiä ja käymään rakentavaa vuoropuhelua tekoälyjärjestelmien tarjoajien kanssa. Tätä tarkoitusta varten Euroopan komission asiantuntijaryhmä¹⁸⁰ laati listan keskeisistä eettisistä ja tekoälyn käytön kannalta oleellisista kysymyksistä. Oheisessa taulukossa on ote niistä:

Ihmisen toimijuus ja ihmisten suorittama valvonta

- 1 Onko opettajan rooli määritelty selkeästi, jotta voidaan varmistaa, että opettaja on aina mukana prosessissa, kun tekoälyjärjestelmää käytetään? Miten tekoälyjärjestelmä vaikuttaa opettajan didaktiseen rooliin?
- 2 Tehdäänkö opiskelijoihin vaikuttavat päätökset niin, että opettajalla on päätöksenteossa aktiivinen rooli, ja voiko opettaja havaita poikkeavuuksia tai mahdollista syrjintää?
- 3 Onko opettajilla ja koulujen johtajilla koulutus ja tiedot, joita tarvitaan järjestelmän tehokkaaseen käyttöön ja sen varmistamiseen, että järjestelmä on turvallinen eikä aiheuta haittaa tai loukkaa opiskelijoiden oikeuksia?

Läpinäkyvyys

- 1 Tietävätkö opettajat ja koulujen johtajat, mitä tekoälyyn perustuvia menetelmiä ja toimintoja järjestelmässä on?
- 2 Ymmärtävätkö opettajat ja koulujen johtajat, miten tekoälyjärjestelmän käyttämät arviointi- tai yksilöllistämisalgoritmit toimivat?
- 3 Ovatko ohjeet ja tiedot saatavilla sekä opettajille että oppijoille selkeässä muodossa?

Monimuotoisuus, syrjimättömyys ja oikeudenmukaisuus

- 1 Onko järjestelmä kaikkien saatavilla samalla tavalla ja esteettömästi?
- 2 Onko käyttöliittymä oppijoiden ikätason kannalta soveltuva ja esteetön? Onko järjestelmän käytettävyyttä ja käyttäjäkokemusta testattu kohdeikäryhmän osalta?
- 3 Onko käytössä menettelyjä, joilla varmistetaan, että tekoälyn käyttö ei johda syrjintään tai epäoikeudenmukaiseen toimintaan kenenkään käyttäjän kohdalla?

Yhteiskunnallinen ja ekologinen hyvinvointi

- 1 Miten tekoälyjärjestelmä vaikuttaa oppijoiden ja opettajien sosiaaliseen ja emotionaaliseen hyvinvointiin?
- 2 Ilmaiseeko tekoälyjärjestelmä selvästi, että sen sosiaalinen vuorovaikutus on simuloitua eikä sillä ole kykyä ymmärtää ja tuntea?
- 3 Käytetäänkö tietoja apuna, kun opettajat ja koulujen johtajat arvioivat opiskelijoiden hyvinvointia, ja jos käytetään, miten tätä seurataan?

Yksityisyyden suoja ja datan hallinta

- 1 Onko olemassa mekanismeja, joilla varmistetaan, että arkaluonteiset tiedot pidetään nimettöminä? Onko käytössä menettelyjä, joilla rajoitetaan pääsy tietoihin vain niitä tarvitseville?
- 2 Onko olemassa mekanismi, jonka avulla opettajat ja koulujen johtajat voivat ilmoittaa yksityisyyteen tai tietosuojaan liittyvistä ongelmista?
- 3 Onko tekoälyjärjestelmä yleisen tietosuojasetuksen mukainen?

Tekninen luotettavuus ja turvallisuus

- 1 Onko käytössä riittävät turvatoimet, joilla järjestelmät suojataan tietoturvaloukkauksilta?
- 2 Onko käytössä strategia, jolla seurataan ja testataan, palveleeko tekoälyjärjestelmä tarkoitustaan ja käytetäänkö sitä suunnitellulla tavalla?
- 3 Onko käytössä asianmukaiset valvontamekanismit tietojen keruuta, tallentamista, käsittelyä, minimointia ja käyttöä varten?

Vastuuvollisuus

- 1 Kuka vastaa tekoälyjärjestelmän tuotosten jatkuvasta seurannasta ja siitä, miten niitä käytetään opetuksen, oppimisen ja arvioinnin tehostamiseen?
- 2 Kuka on vastuussa tekoälyjärjestelmän hankintaa ja toteutusta koskevista lopullisista päätöksistä?
- 3 Onko olemassa palvelutasosopimus, jossa esitetään selkeästi tuki- ja ylläpitopalvelut sekä toimet, joihin on ryhdyttävä ilmoitettujen ongelmien ratkaisemiseksi?

¹⁸⁰ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

Kirjoittajien esittelyt

FM **Pipsa Havula** on Faktabaarin vastaava toimittaja, faktantarkistaja ja freelance-toimittaja. Faktabaariin helmikuussa 2022 mukaan tullut Havula on perehtynyt erityisesti erilaisten kuva- ja videohujauksen tunnistamiseen, faktantarkistukseen sekä digitaalisen informaatiolukutaidon ja faktantarkistustaitojen opettamiseen koulutusten ja videotuotannon kautta.

Tiina Härkönen toimii johtavana asiantuntijana Sitran Demokratia ja osallisuus -teemalla, Digitaalinen valta ja demokratia -projektissa. Hän on tehnyt pitkän uran yritysmaailmassa datan ja tietoverkkojen parissa, markkinoinnin, viestinnän ja liiketoiminnan kehittämisen johtotehtävissä.

Harto Pönkä (KM) on perehtynyt monipuolisesti mm. verkko-opetuksen pedagogiikkaan, mediakasvatukseen, sosiaaliseen mediaan ja tietosuojaan. Hän on toiminut kouluttajana vuodesta 2008 lähtien sekä julkaissut sosiaaliseen mediaan liittyviä kirjoja ja artikkeleja. Pönkä tekee koulutuksia ja analyyssejä yrityksille, yhdistyksille ja julkishallinnon organisaatioille. Pönkä työskentelee yrityksissään Innowisessa ja Tweepsissä.

Minna Aslama Horowitz on Helsingin yliopiston viestinnän oppiaineen dosentti, tutkija Strategisen tutkimuksen neuvoston DECA-hankkeessa ja Faktabaarin viestintäpolitiikan asiantuntija EU/EDMO-NORDIS-verkostossa.

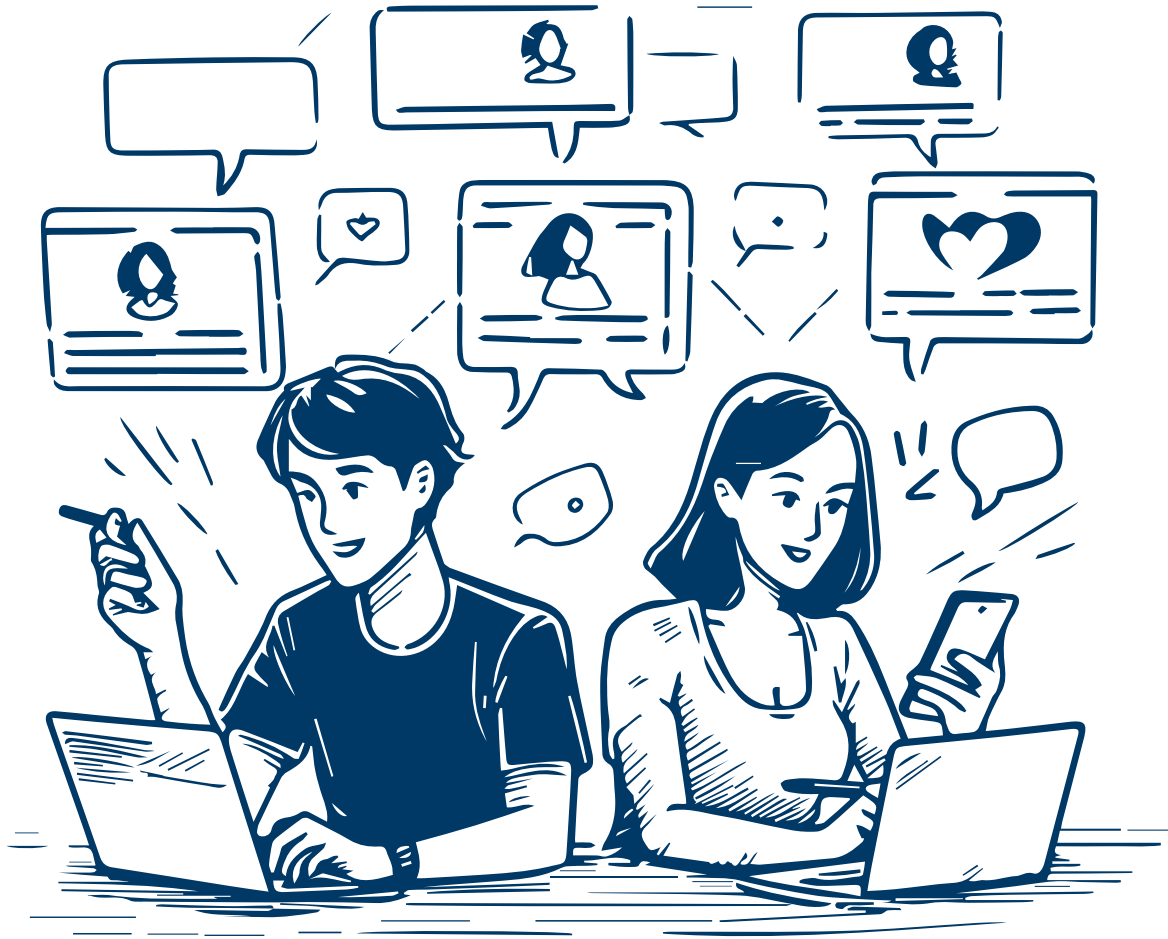
Aki Saariaho on opettanut Espoossa yli 20 vuotta englantia, filosofiaa ja elämäntutkimustietoa. Aki kiinnostaa oppiminen ja sivistys - ennen kaikkea miksi ja miten opitaan sekä mitä sivistys on tänä päivänä. Perheen parissa ja pelaamisen eri muodoissa vietetyn

ajan lisäksi hän harrastaa yhdistyksiä, muun muassa näitä: POE:n hallituksen jäsen (OAJ:n pedagogisten yhdistysten liitto), Feto ry:n hallituksen varajäsen (filosofian ja elämäntutkimustiedon opettajien yhdistys), Systems Change Finland ry:n hallituksen jäsen sekä eOppimiskeskuksen hallituksen varajäsen.

FT **Kari Kivinen** on Faktabaarin digitaalisesta informaatiolukutaidosta vastaava koulutusalan asiantuntija. Hänellä on toiminut opettajana, rehtorina ja pääsihteerinä kansainvälisissä kouluissa ja tehtävissä Suomessa, Luxemburgissa, Belgiassa ja Espanjassa. Hän on mukana Euroopan komission digitaalisen osaamisen asiantuntijaryhmissä.

Tapani Saarinen (Ph.D.) on toiminut luokanopettaja Eurooppa-koulussa Brysselissä sekä Helsingin yliopiston Viikin normaalikoulussa. Hän on ollut mukana asiantuntijaryhmissä, jotka avustivat Euroopan komissiota tekoälyn opetuskäytön eettisten ohjeiden laadinnassa sekä digitaalisen oppimateriaalin laadun selvitystyössä. Saarinen on osallistunut keskusteluun tekoälyn opetuskäytöstä muun muassa Euroopan Unionin ja European Schoolnetin järjestämissä webinaareissa. Teknologiateollisuuden 100-vuotissäätiö palkitsi Saarisen innovatiivisesta pedagogisesta työstä vuonna 2017.

Matt Cronin on toiminut Valkoisen talon kansallisen kyberturvallisuusosaston (ONCD) johtajana. Cronin nimitettiin tehtävään Yhdysvaltain oikeusministeriöstä, jossa hän toimii kansallisen turvallisuuden ja kyberrikollisuuden koordinaattorina. Cronin on myös Fulbright-tutkija, joka tutkii tekoälyn ja muiden uusien teknologioiden poliittisia vaikutuksia.



FaktaBaari **EDU**

[Faktabaari.fi/dil/ai](https://faktabaari.fi/dil/ai)