

HUOMIO



EI YKSIN RIITÄ
HYBRIDIÄLY OPETUKSESSA

Centria. Oppimateriaaleja, 7.

Emilia Nygård

HUOMIO EI YKSIN RIITÄ - HYBRIDIÄLY OPETUKSESSA

Centria-ammattikorkeakoulu, 2023.

Kokkola



Euroopan unioni
Euroopan sosiaalirahasto

Vipuvoimaa
EU:lta
2014–2020



Elinkeino-, liikenne- ja
ympäristökeskus

centria
ammattikorkeakoulu

modo
Modulaarinen, digitaalinen
ja älykäs oppimisympäristö

Hanketta rahoitetaan Euroopan sosiaalirahastosta osana Euroopan unionin COVID-19 -pandemian johdosta toteuttamia toimia.

Julkaisija: Centria-ammattikorkeakoulu, Talonpojankatu 2, 67100 Kokkola
Tekstit ja taitto: Emilia Nygård, TKI-kehittäjä, Centria-ammattikorkeakoulu
Tekoälyä (ChatGPT, Bing Chat) on käytetty käännösten apuna
Kansi: Bing Image Creator

Centria. Oppimateriaaleja, 7
ISBN: 978-952-7173-79-4
ISSN 2342-9348

SISÄLLYS

JOHDANTO	6
CASE GITHUB COPILOT	8
MIKÄ TEKOÄLY?.....	10
Koneoppiminen.....	13
Neuroverkot ja syväoppiminen	14
GENERATIIVINEN TEKOÄLY.....	18
Perusmalli ja kielimalli	18
Transformer-arkkitehtuuri	21
Miten kaikki tapahtuu?	24
Tekoälytaide.....	26
Teknologiajättien kilpajuoksu	28
LAINSÄÄDÄNTÖ	34
TEKOÄLYN ETTISIÄ HAASTEITA	38
Data	40
Oikeudenmukaisuus ja vinoumat	41
Hinta ja vallan keskittyminen	42
Ympäristövaikutukset	43
Tekijänoikeus, yksityisyys ja turvallisuus	44
Polarisoituminen ja vaikutukset demokratiaan	45
Vaikutus työvoimaan	45
TEKOÄLY OPETUKSESSA.....	49
LÄHTEET	56

**“If you missed the 90s –
WWW, Mosaic, and the
founding of Amazon and
Google – you don’t want
to miss this opportunity.”**

Amazon

Työnteon tuottavuutta on kautta historian yritetty tehostaa. 1700– ja 1800–luvun taitteessa höyrykone mullisti teollisia tuotantomenetelmiä. Tämän ensimmäiseksi teolliseksi vallankumoukseksi nimetyn ajan myötä toistuvia työtehtäviä ryhdyttiin korvaamaan mekaanisilla koneilla. 1800–luvun lopussa keksittiin sähkö- ja polttomoottorit, joiden myötä toinen teollinen vallankumous laajensi teollista toimintaa. 1900–luvun puolivälin tienoilla alkaneen kolmannen teollisen vallankumouksen teknisenä innovaationa pidetään mikropiiriä – tietokoneet tulivat osaksi yhteiskuntaa. Elämme nyt neljännen teollisen vallankumouksen aikaa. Digitalisaatiokehitys jatkuu, kun fyysinen ja digitaalinen maailma tulevat entistä lähemmäs toisiaan. Ajanjakso tunnetaan myös nimellä Teollisuus 4.0 (Industry 4.0) ja sitä voidaanakin kutsua tekoälyllä tehostetuksi tuottavuusvallankumoukseksi.

Uuden teknologian sääntely on tarpeen, mutta siihen mitä ja miten sitä säännellään ei ole helppoa ratkaisua. Karkeasti katsottuna, kaikkialla ollaan yhtä mieltä siitä, että tekoälyn käytön on oltava avointa ja turvallista ja tästä lähtökohdasta EU:ssa valmistellaan lakeja minimoimaan tekoölyyn liittyviä riskejä ja varmistetaan sen eettinen käyttö, niin pitkälle kuin se suinkin on mahdollista. Nopeasti kehittyvä teknologia yhdistettynä hitaasti etenevään lakipykälän valmisteluun aiheuttaa haasteita muiden muassa teknologian kehittäjille: kun pelikentän säännöt ovat vasta muodostumassa, liittyy investointeihin entistä suurempia riskejä.

EU säätelee proaktiivisesti tekoälyjärjestelmiä asettaen etusijalle turvallisuuden, läpinäkyvyyden ja syrjimättömyyden. Ehdotettu lainsäädäntö luokittelee tekoälyjärjestelmät eri riskitasoihin, asettaa vaatimuksia korkean riskin tekoälyjärjestelmille, edellyttää läpinäkyvyyttä sekä valvomista lain täytäntöönpanoon. EU:n tekoälyasetuksesta on tulossa maailman ensimmäinen, kattava tekoälylainsäädäntö. OpenAI:n Sam Altman on ilmaissut huolensa tekoälyn nopeaan kehitykseen liittyen ja painottanut maailmanlaajuisen sääntelyn tarpeellisuutta. Toukokuussa 2023 Altman kuitenkin totesi, että mikäli EU:ssa säädettävät lait rajoittavat yrityksen toimintaa liikaa, OpenAI saattaa jättää EU:n markkinat kokonaan. Tekoälykilpailu käy kuumana, kun päivittäin satoja uusia tekoälytyökaluja tuodaan markkinoille. Kuten Stanfordin yliopiston Cyber Policy Centerin kansainvälisten asioiden päällikkö Marietje Schaakekin on todennut, olisi viisasta pitää toimitusjohtajat poissa tekoälyn sääntelystä.

Uusien teknologioiden myötä ihmiset ovat aina olleet huolissaan sen vaikutuksista ja etenkin työpaikkojen kohtalosta. Esimerkiksi maatalouskoneet, kirjapaino, autot ja internet ovat kaikki parantaneet dramaattisesti elämänlaatua, mutta samalla myös luoneet paljon uusia työpaikkoja. Sama tulee luultavasti pätemään myös tekoälyyn. Meidän tulee sopeutua ja opetella elämään sen kanssa, koska se on tullut jäädäkseen.

Ammattikorkeakoulujen rehtorineuvosto Arene ry julkaisi toukokuussa 2023 omat suosituksensa tekoälyn hyödyntämisestä ammattikorkeakouluissa. Tekoäly voidaan laskea kuuluvaksi digitaaliseen kompetenssiin ollen näin osa työelämävalmiuksia. Ammattikorkeakoulujen tehtävä on varmistaa, että työelämätaidot valmistuvilla oppilailta ovat riittävällä tasolla. Arene suosittelee ammattikorkeakouluja valitsemaan mahdollisimman suosivan linjan tekoälytyökalujen käytössä, mutta varmistamaan niiden tarkoituksenmukaisen ja eettinen käytön sekä henkilökunnan ja opiskelijoiden valmiudet hyödyntää tekoälyä vastuullisesti. (Arene, 2023.)

MODO (=modulaarinen, digitaalinen ja älykäs oppimisympäristö) -hankkeessa on kehitetty opintojaksopilotteja, joiden tavoitteena on kehittää toisen asteen opiskelijoiden digitaalisia taitoja, motivaatiota ja yleisiä osaamisvaatimuksia. Piloteissa on testattu Centria-ammattikorkeakoulun käyttämän itsLearning-oppimisympäristön aktiviteetteja ja rakenteita. Toisen ja korkea-asteen opetushenkilöstölle on järjestetty digitaalisia taitoja edistäviä digityöpajoja, joiden myötä opettajat ovat saaneet uusia työkaluja käyttöönsä opintojaksojen monipuolisuutta ja mielekkyyttä lisätäkseen. Hankeen rahoittajana toimii Euroopan sosiaalirahasto (ESR). (Rakennerahastotietopalvelu; Murto, M, 2023.) Esimerkiksi juuri tekoälyn avulla opintojaksojen monipuolisuutta voidaan lisätä.

Yksi merkittävä tekoälyn käyttöönottoon liittyvistä haasteista on todettu olevan se, ettei osapuolilla ole riittävä ymmärrystä siitä, miten tekoäly todellisuudessa toimii. Tämä kävi ilmi myös Digivisio2030-hankkeen selvityksessä "Tekoälypohjainen oppijan ohjaus ja neuvonta". (Hotakainen & Kivinen, 2023). Tämän materiaalipaketin tarkoitus on parantaa opettajien tietoutta tekoälystä, jotta sen käyttö ja soveltaminen opetuksessa olisi sujuvampaa.

Kun GitHub julkaisi tekoälytyökalu Copilotin vuonna 2021 ajatuksena oli, että se parantaisi ohjelmistokehittäjien tuottavuutta. Copilot on GitHubin ja OpenAI:n yhteistyö; se on tekoälyllä toimiva työkalu, joka on integroitu GitHubin koodivarastoon ja jota voidaan käyttää koodin luomiseen, parannusehdotuksiin ja kontekstikohtaisten ehdotusten antamiseen. Vuonna 2022 GitHub teetti tutkimuksen, johon osallistui yli 2000 amerikkalaista ohjelmistokehittäjää. Näistä noin 60 % teki ohjelmistokehittäjän työtä ammatikseen, loput olivat alan opiskelijoita tai harrastelijoita. Tutkimuksen tuloksena valtaosa, yli 90 % vastaajista oli sitä mieltä, että GitHub Copilot auttoi heitä suoriutumaan tehtävistään nopeammin. (Kalliamvakou, 2022.)

Tutkimuksen avulla haluttiin selvittää, kokivatko GitHub Copilotia käyttävät ohjelmistokehittäjät hyötyvänsä työkalusta muullakin tavoin kuin ajallisesti. 60–75 % käyttäjistä raportoivat olevansa tyytyväisempiä työhönsä ja tunsivat olonsa vähemmän turhautuneiksi. Lisäksi he kertoivat voivansa käyttää aikansa mielekkäämpiin työtehtäviin käyttäessään GitHub Copilotia. 75 % vastaajista oli sitä mieltä, että GitHub Copilot helpotti heitä pysymään työn imussa ja 87 % kertoi pystyvänsä GitHub Copilotin avulla säilyttämään henkisen panoksensa työhön etenkin toistuvissa tehtävissä. Suurimmat kyselyn oivallukset liittyivät nimenomaan siihen, että antamalla GitHub Copilotin tehdä ohjelmistokehittäjän ”tylsät” ja toistuvat työt, vähentää se työn kognitiivista rasitusta. Tämä puolestaan vapautti kehittäjien aikaa tehdä mielestään merkityksellisempää työtä: työtä, joka vaatii monimutkaista, kriittistä ajattelua ja ongelmanratkaisutaitoja. Kaikki tämä puolestaan johti onnellisempaan ja tyytyväisempään työntekijään. (Kalliamvakou, 2022.)

Tutkimuksen pohjalta GitHub rekrytoi erilliseen tutkimukseen 95 ohjelmistokehittäjää ja jakoi heidät satunnaisesti kahteen ryhmään. Testihenkilöiden tehtävänä oli kirjoittaa HTTP-palvelin JavaScriptillä. Toinen ryhmä (45 henkilöä) käytti apunaan GitHub Copilotia, jota toinen ryhmä (50

henkilöä) ei käyttänyt. GitHub Copilotia käyttävät suoriutuivat tehtävästä yli tuplasti verkkiryhmäänsä nopeammin. (Kalliamvakou, 2022.)

Kuten GitHubin Copilot, monen muun tekoälytyökalun hienous piilee juuri tehokkuuden parantamisessa. Tehokkuudella ei tarkoiteta välttämättä vain nopeutta suoritua jostakin tehtävistä, vaan se voi ilmetä eri tehtävissä eri tavoin. Tekoälytyökalujen avulla voimme vapauttaa resursseja. Kun saamme keskittyä tekemään itsellemme merkityksellistä ja mielekästä työtä, voimme myös henkisesti paremmin. Eikö tämän pitäisi olla kaikkien työnantajien yksi päätavoitteista – tehokas ja hyvinvoiva henkilöstö?

Maailman talousfoorumin Future of Jobs-raportti on kolmen vuoden välein julkaistava tutkimus. Vuonna 2023 julkaistun tutkimuksen tuloksia leimaa ajan epävarmuus ja nopea teknologiatekijäkehitys. Tutkimus osoitti että 23 % työtehtävistä tulee todennäköisesti muuttumaan. Tutkimuksen mukaan eniten kasvua, määrällisesti työpaikkoja, tulee olemaan koulutus- ja maataloussektorilla. Koulutussektori tulee kasvamaan arvioiden mukaan noin kymmenen prosenttia, johtaen yli 3 miljoonaan uuteen ammatilliseen ja yliopisto- ja korkeakouluopettajan työpaikkaan. (World Economic Forum, 2023d.)

Tutkimuksessa selvitettiin myös, millaisia taitoja yritykset työntekijöissään arvostavat. Tällaisia olivat muun muassa itseluottamus, joustavuus, ketteruus, motivaatio ja uteliaisuus. Samalla kun yritysten tulee uudistaa järjestelmiään ja prosessejaan vastaamaan tätä päivää on syntynyt tarve aivan uudelle osaamiselle: työmarkkinat tarvitsevat ihmisiä, jotka pystyvät muutoksen mukana ja pystyvät johtamaan itseään. Varmuudella ei kuitenkaan voida sanoa, millaisia kykyjä ja taitoja tulevaisuudessa työntekijöiltä tarvitaan, mutta tiedetään että tulevaisuuden taidot liittyvät vahvasti muutokseen ja siihen adaptoitumiseen. (World Economic Forum 2023d.) Pystyykö koulutussektori nykyisellään vastaamaan tähän työelämän tarpeeseen?

työnantajien arvostamia
työelämätaitoja

MIKÄ TEKOÄLY?

ani, agi, asi

Tekoäly, AI (Artificial Intelligence), määritellään usein järjestelmäksi, joka voi suorittaa sellaisia tehtäviä, esimerkiksi ongelmanratkaisua, päätöksentekoa, oppimista ja kielten ymmärtämistä, musiikin säveltämistä, auton ohjaamista tai esineiden ja asioiden tunnistamista kuvista, jotka yleensä edellyttäisivät inhimillistä älykkyyttä ja aivojen prosessointikykyä. Tekoäly on tieteenala: se on joukko käsitteitä ja ongelmia sekä menetelmiä niiden ratkaisemiseksi. Autonomisuus ja adaptiivisuus ovat tekoälylle tyypillisiä ominaisuuksia: se suorituu tehtävistä ilman jatkuvaa opastusta ja sillä on kyky oppia kokemastaan ja parantaa suorituskyykyään oppimansa perusteella. Artificial intelligence-termin, takana uskotaan olleen yhdysvaltalainen tietojenkäsittelytieteilijä John McCarthy (1927–2011). Häntä pidetään myös tekoälytieteen isänä. (Elements of AI, 2018a&b.)

Tekoäly voidaan jakaa kapeaan- ja yleiseen tekoölyyn, joista kapealla (tai heikolla) tekoölyllä (ANI, artificial narrow intelligence) viitataan sellaisiin tekoälyjärjestelmiin, jotka on suunniteltu suorittamaan tiettyjä yksittäisiä tehtäviä, kuten kuvien tunnistaminen tai tekstin kääntäminen. Ne voivat suorittaa älykkäitä toimintoja, mutta eivät ole älykkäitä. Yleinen (tai vahva) tekoäly (AGI, artificial general intelligence) puolestaan viittaa aidosti älylliseen ja tietoiseen järjestelmään. Yleinen tekoäly on suurilta osin teoreettinen käsite, eikä sitä käytännössä ole olemassa. Se viittaa inhimilliseen älykkyyteen verrattavissa oleviin tekoälyjärjestelmiin, joilla on kyky ymmärtää, oppia ja soveltaa tietoa. (Elements of AI, 2018a&b.) Mikäli yleinen tekoäly pystyttäisiin kehittämään, johtaisi se teoriassa nopeasti supertekoölyyn (ASI, artificial super intelligence). ASI viittaa järjestelmään, joka pystyy tuottamaan asioita, joita ihminen ei ymmärrä.

Tekoälystä puhuttaessa nousee usein esiin Turingin testi. Se on Alan Turingin (1924–1954) kehittämä testi, jonka ajatuksena on, että olennon on oltava älykäs, jos ihminen ei pysty erottamaan sitä sen käyttäytymisen perusteella toisesta älykkästä olennosta. Eli jos ihminen ei huomaa keskustelewansa koneen kanssa, on kone tällöin älykäs. Testissä ihmishaastattelija on vuorovaikutuksessa kahden pelaajan kanssa viestien välityksellä. Ihmishaastattelijan tehtävä on yrittää erottaa, kumpi pelaajista on kone ja kumpi ihminen. Mikäli haastattelija ei kykene erottamaan konetta ihmisestä, sanotaan koneen läpäisseen testin. Testiä on kritisoitu siitä, että se ei välttämättä mittaakaan älykkyyttä vaan oikeammin ihmismäisyyttä. Yksi argumentoijista on yhdysvaltalainen filosofi John Searle (1932–) joka kyseenalaisti sen, onko älykkyys ja älykäs toiminta sama asia. Searlen argumentin mukaan kone voi näyttää toimivan älykkäästi ja esimerkiksi läpäistä Turingin testin, mutta se ei tarkoita, että se olisi älykäs.

Tekoälyn jaottelu kapeaan ja yleiseen (tai heikkoon ja vahvaan) perustuu nimenomaan erotteluun älykkyyden ja älykkään toiminnan välillä. (Elements of AI, 2018a.)

On vaikeaa vetää rajaa sille, mikä määritellään tekoälyksi ja mikä ei. Tekoälyä on kaikkialla, emmekä välttämättä miellä tekoälyksi enää sellaisia asioita, jotka ovat ottaneet pysyvän roolin arjessamme. Tekoäly suorittaa toimintoja, joita emme tyypillisesti yhdistä siihen, kuten kasvotunnistus, koneellinen kielenkääntäminen, henkilökohtaiset uutissyötteet, ehdotukset ja mainokset esimerkiksi sosiaalisen median kanavissa tai reittisuunnittelu (esim. Google Maps). Hakukoneetkin käyttävät tekoälyä parantamaan hakutulosten tarkkuutta ja osuvuutta. (World Economic Forum, 2023b.)

Tekoälyn määritelmä täsmentyy ja täydentyy käsi kädessä teknologian kehityksen kanssa. Koska tekoäly on kapea-alaista sen kyky ratkaista yhtä ongelmaa ei kerro mitään sen kyvystä ratkoa muita ongelmia. Tästä syystä tekoälyjärjestelmiä ei ole mielekästä verrata keskenään niiden älykkyyden perusteella. On turhan jyrkkää vetää kyllä/ei-linjauksia siitä, onko jokin tekoäly vai ei. Tarkoituksenmukaisempaa olisi tarkastella tekoälyä ominaisuutena, jonka määrä voi vaihdella. (Elements of AI, 2018a.)

”riipaus tekoälyä”

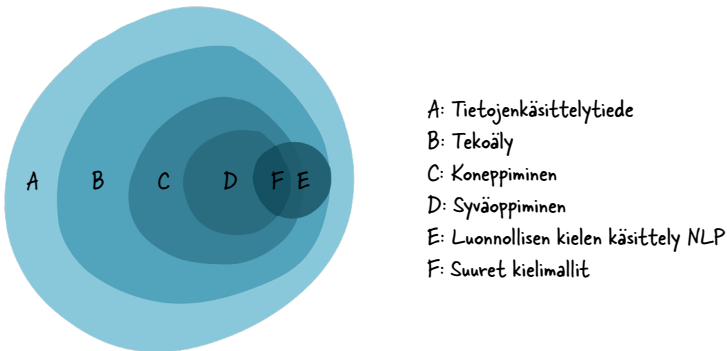
Tekoälyn tavoitteena on varmistaa, että sen tuottamat toimet vastaavat meidän sille asettamia tavoitteita. Sen perimmäinen tavoite on pystyä suoriutumaan tehtävistä kuten ihminen, tai jopa paremmin. Tulee muistaa, että tekoäly on teknologia: se ei ole itsessään hyvä tai paha, vaan sen hyvyys tai pahuus on meistä kiinni. Tekoäly on joukko algoritmeja, joiden mukaan tietokone toimii. Ei ole kuitenkaan järkevää luottaa siihen, että algoritmien vaikutus, ei vaan tekoälyssä vaan algoritmien vaikutus ylipäättään, olisi välttämättä aina positiivinen. (World Economic Forum, 2023b.)

Yleinen tekoäly ei tule valloittamaan maailmaa yhdessä yössä, vaan sen vaikutukset tulevat vähitellen kasvamaan. Jokaisen tekoälyn ottaman edistysaskeleen myötä sillä tehtyjen tehtävien määrä nousee merkittävästi. Yleisen tekoälyn valloitusta on odotettu jo pitkään ja sen on jo monesti sanottu olevan ”aivan nurkan takana”. Jotkut asiantuntijat arvioivat yleisen tekoälyn olevan käytössä vuosisadan loppuun mennessä, toisten mielestä sitä ei tule ikinä tapahtumaan. Elämme aikaa, jossa saatetaan ajatella tekoälyn tuovan ratkaisun jokaiseen elämän osa-alueeseen. Suuressa mittakaavassa näin tulee olemaan, mutta tulee myös aikakausi, jolloin huomaamme, ettei se pystynytäkään kaikkeen, mihin kuvittelimme sen pystyvän. (World Economic Forum, 2023a&b.)

**“The good news
about computers is
that they do what
you tell them to do.
The bad news is
that they do what
you tell them to
do.”**

Ted Nelson, philosopher, computer
scientist and sociologist

Tekoäly kuuluu tietojenkäsittelytieteen piiriin. Kuten aiemmin mainittiin, autonomisuus ja adaptiivisuus ovat tekoälylle tyypillisiä ominaisuuksia. Tekoäly on yleinen termi, joka kuvaa ohjelmistoa, joka voi älykkäästi työskennellä kohti tiettyä tehtävää. Koneoppimisen, joka on yksi tekoälyn osa-alue, avulla tekoälysovelluksista saadaan mukautuvia eli adaptiivisia. Koneoppimisen avulla järjestelmät parantavat suorituskykyään kokemuksen ja datamäärän karttuessa – näinhän me ihmisetkin toimimme. Syväoppiminen taas on koneoppimisen osa-alue. (Elements of AI, 2018a.)



KUVA 1. Tietojenkäsittelytiede (Elements of AI, 2018a mukailleen.)

On tilanteita, joissa me ihmiset osaisimme mahdollisesti ennustaa algoritmia paremmin, mutta koneoppimisalgoritmi voi kuitenkin olla ihmistä käytännöllisempi sen nopeuden ja tehokkuuden takia. Algoritmien ja menetelmien käyttökohteet vaihtelevat ja onkin hyvä korostaa, että on tärkeää löytää tiettyyn tilanteeseen sopivin tai sopivimmat menetelmät. Virheiltilta on mahdotonta välttyä ja tilanteesta riippuen ”hyvä” on riittävä tarkkuus, eikä täydellisyyteen ole tarkoituksenmukaista pyrkiäkään. Tämä toki riippuu sovelluksesta: itseohjautuvan auton kanssa täydellisyys on tavoiteltavaa, kun taas Netflixin suositusten kanssa ei niin haittaa, jos jokin suositus ei juuri osu omaan makuusi. (Elements of AI, 2018c.)

Koneoppiminen voidaan jakaa karkeasti kolmeen eri ryhmään: ohjattu ja ohjaamaton koneoppiminen sekä vahvistusoppiminen. Ohjatussa koneoppimisessa koneen tarkoitus on tunnistaa ja lajitella data oikeaan luokkaan. Opetusdataa apuna käyttäen järjestelmä opetetaan (ts. koulutetaan) tunnistamaan oikea vastaus. Opetusdata on käytännössä joukko esimerkkejä, joiden jokaisen kohdalle on merkitty oikea vastaus. Se sisältää siis sekä syöte- että niitä vastaavat tulostiedot. Merkityn opetusdatan myötä kone tai algoritmi oppii itsenäisesti ennustamaan oikeat vastaukset mille tahansa syönteelle. Mallin ennustustarkkuutta arvioidaan opetusprosessin jälkeen testidatan avulla. Menetelmiä ohjattuun koneoppimiseen on lukuisia. Kaikissa perusperiaatteena on kuitenkin löytää datasta jokin säännönmukaisuus ja käyttää sitä oikean vastauksen ennustamiseen. Ohjattua koneoppimista käytetään esimerkiksi käyttäjän aiemman verkko-käyttäytymisen perusteella ennustaminen. (Elements of AI, 2018c.)

ohjattu oppiminen:
 säännönmukaisuuden löytäminen

ohjaamaton oppiminen:
ei valmiita luokkia

Ohjaamattomassa koneoppimisessa ei ole valmiiksi määriteltyjä luokkia, kuten ohjatussa koneoppimisessa, vaan tehtävänä on löytää datasta jokin siinä toistuva ja tunnistettava rakenne. Datassa on syötetiedot, mutta sitä vastaavaa tulostietoa ei joko ole, tai sitä ei hyödynnetä. Ohjaamatonta oppimista on esimerkiksi datan visualisointi, jossa samankaltaiset esimerkit sijoitetaan kuvaan lähelle toisiaan ja vastaavasti erilaiset esimerkit kauas toisistaan. Niin sanotussa ryhmittelyssä (engl. clustering), ryhmiä muodostetaan siten, että ryhmään kuuluvat esimerkit ovat keskenään samanlaisia, mutta eroavat muiden ryhmien esimerkeistä. Esimerkiksi ruokaostosten datan analysoinnissa käytetään ryhmittelyä ja sen tuloksena samankaltaisia tuotteita ostavat asiakkaat ovat visualisoinnissa lähempänä toisiaan. (Elements of AI, 2018c.)

vahvistusoppiminen:
tavoitetilan saavuttaminen

Vahvistusoppimisessa pyritään pääsemään johonkin tavoitetilään, muuttamalla eri toiminnoilla järjestelmän tilaa. Järjestelmä saa palautetta siitä, kuinka hyvin tehtävät auttavat sitä saavuttamaan tavoitetilan, mutta tieto siitä, toimiko se oikein vai väärin tulee mahdollisesti viiveellä. Vahvistusoppimista voidaan käyttää silloin, kun pystymme määrittelemään, kuinka konetta palkitaan ja seuraamaan sen toimintaa. Itseohjautuva auto on esimerkki vahvistusoppimisesta. Näiden kolmen koneoppimiskategorian rajat ovat häilyviä. Niitä myös yhdistellään muodostaen uusia kategorioita, kuten puoliohjattu koneoppiminen. (Elements of AI, 2018c.)

Kun kone tai algoritmi parantaa suoritustaan vähitellen tekoälyongelman suhteen, sen ajatellaan olevan oppiva. Ohjatussa oppimisessa kone opetetaan yhdistämään kerätty data sitä vastaavaan taustatietoon. Ohjaamattomassa oppimisessa taustatieto puuttuu, mutta halutaan oppia ymmärtämään datan rakennetta esimerkiksi, että ryhmittyykö data jollakin tavoin. Vahvistusoppimisessa kone muuttaa toimintaansa saadun palautteen perusteella. (Hursti ym. 2023a.)

Neuroverkot ja syväoppiminen

neuroni: $f(x)=3x+5$

Neuroverkoista puhuttaessa voidaan viitata joko biologisiin neuroverkkoihin tai tietokoneilla simuloitaviin, keinotekoisin neuroverkkoihin (ANN, artificial neural network). Kummasta tahansa puhutaan, koostuu se valtavasta määrästä yksittäisiä yksiköitä eli neuroneja, joiden tehtävä on lähettää ja vastaanottaa viestejä toisilta neuroneilta. Keinotekoinen neuroverkko on eräänlainen koneoppimismalli, jossa neuronit ovat matemaattisia funktioita. Jokainen neuroni on siis yksinkertainen matemaattinen funktio, joka laskee tulosteen syötteiden perusteella. (Elements of AI, 2018e; Muehmel, 2023.)

neuroverkko:



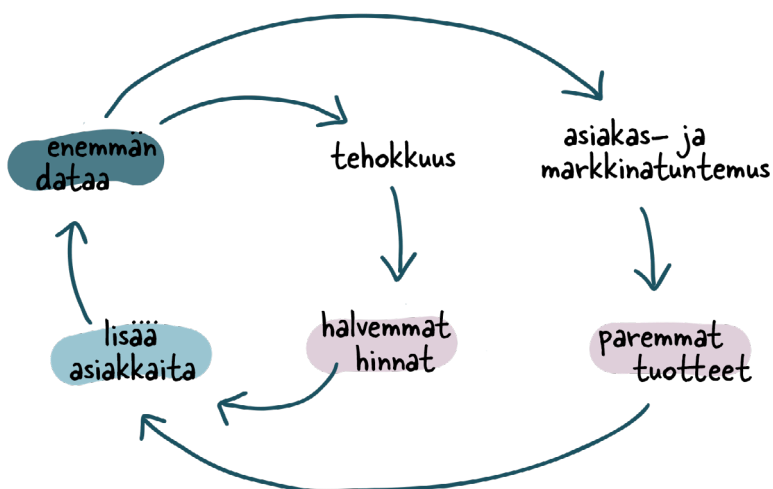
Neuroverkossa neuronit on usein jaettu moneen kerrokseen. Tällaista koneoppimisen kerroksellista menetelmää kutsutaan syväoppimiseksi. Mitä useampi kerros, sitä syvämpi verkko. Mitä syvämpi verkko, sitä vaikeampi opetusprosessi ja sitä enemmän dataa ja laskentatehoa tarvitaan. Onnistuessaan tämä mahdollistaa kuitenkin hyvinkin monimutkaisten ilmiöiden oppimisen. (Elements of AI, 2018e.)

Yksittäinen neuroni ei pysty suorittamaan kuin hyvin yksinkertaista tehtävää, mutta yhdessä neuronit voivat suorittaa hyvinkin monimutkaisia toimintoja. Neuroverkko voi prosessoida valtavaa määrää tietoa samanaikaisesti,

koska jokainen neuronin voi prosessoida tietoa toisesta neuronista riippumatta. Neuroverkon voima piilee nimenomaan neuroneiden välisissä yhteyksissä. Jokainen neuronin on yhteydessä toisiin neuroneihin, ja jokaisen yhteyden vahvuus määritellään numeerisella painolla. Nämä painot määräävät, kuinka paljon yhden neuronin tulostetta otetaan huomioon seuraavan neuronin syötteenä. (Muehmel, 2023.)

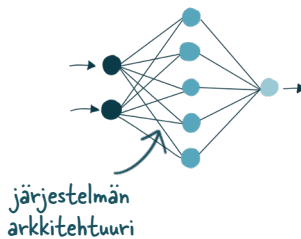
Neuroneissa on siis joukko mukautuvia painoja. Asian havainnollistamiseksi, kauppalaskussa painot vastaisivat ostettujen tuotteiden hintoja. Painot opitaan lähes aina datasta. Neuroverkko oppii, kun painojen arvoja säädellään niin, että saadut tulokset ovat oikeita. Jo pieni muutos painojen arvoihin, muuttaa lopputulosta. Kauppalaskutapauksessa tuotteiden hintoja (painoja) siis säädeltäisiin niin, että niiden yhteenlaskettu määrä saataisiin vastaamaan laskun loppusummaa. Painojen määrä voi kasvaa neuroverkoissa jopa sadoiksi miljardeiksi ja tällaisen määrän optimointi vaatii valtavan määrän dataa ja laskentatehoa. (Elements of AI, 2018e.)

Tietokoneiden laskentatehojen moninkertaistuminen ja datan tallentamisen edullisuus ovat osasyitä tekoälyn tämänhetkisellem laajuudelle. Myös suurien data-aineistojen kerääminen on tullut mahdolliseksi tietoverkkojen ja internetin myötä. Ilman näitä seikkoja tekoäly ei nykyisessä mittakaavassa olisi mahdollinen. (Chui, Roberts, Rodchenko, Singla, Sukharevsky, Yee & Zurkiya, 2023.). Datamäärä on viime vuosikymmenien aikana kasvanut ja sen hinta on laskenut, koska sen säilyttäminen suuressa mittakaavassa on tullut mahdolliseksi ja edulliseksi. Monet yritykset ovat hyödyntäneet ja keränneet dataa eri tarkoituksiin. Tekoälyn kehittymisen myötä data on noussut uuteen arvoon ja tämä hyödyttää erityisesti niitä yrityksiä, joilla on paljon omaa dataa. (Thomas, 2023.)



KUVA 2. Data-oppimissykli. (Thomas, 2023 mukailten.)

Kyky käsitellä ristiriitaista ja epävarmaa informaatiota on osasy sille, miksi nykyaikaiset tekoälymenetelmät toimivat käytännössä. Epävarmuutta voi mitata ja käsitellä numeerisesti. Kuuluisan testinsä lisäksi Alan Turing oivalsi, että kaikki mikä voidaan suorittaa laskennallisesti, on automatisoitavissa. Lähes kaikki tekoälysovellukset käyttävät epävarmuuden hallintaan todennäköisyyksiä. Todennäköisyyksiä apuna käyttäen päättelyä voidaan siis automatisoida. (Elements of AI, 2018b&d.)



Se kuinka järjestelmä toimii, riippuu siitä, kuinka neuronit on kytketty toisiinsa eli järjestelmän arkkitehtuurista. Hyvin yksinkertaisessa neuroverkko-arkkitehtuurissa jokainen neuronin voi olla kytketty sen yläpuolella olevaan neuroniin. Toisissa tapauksissa neuronin voi olla kytketty ainoastaan joihinkin toisiin neuroneihin sen lähellä. Tällaista menetelmää hyödynnetään konvoluutioverkoissa, joista on tullut kuluneen vuosikymmenen aikana kuvantunnistuksen perusta. (Muehmel, 2023.)

Jos datasta halutaan oppia tunnistamaan kuvapiirteitä, voidaan neuroverkkoon sisällyttää tällainen konvoluutiokerros. Neuroverkkoa, joissa on konvoluutiokerros, kutsutaan konvoluutioneuroverkoksi. Konvoluutioverkot voivat tunnistaa kohteita riippumatta siitä missä kohde kuvassa sijaitsee. Aikana ennen konvoluutioverkkoja, jos olisimme halunneet opettaa verkon tunnistamaan kissan korvan sekä kuvan keskeltä että alareunasta, meidän olisi pitänyt syöttää sille sellaista opetusdataa, jossa kissan korva olisi nimenomaan näissä paikoissa. Jos kissan korva sen sijaan olisi sijainnut kuvan yläreunassa verkko ei olisi tunnistanut sitä kissan korvaksi, ellei sille olisi opetusdatassa sellaista opetettu. Konvoluutioverkon myötä tarvittavan datan määrä tippui murto-osaan. (Elements of AI, 2018e.)

Verkon alimmilla tasoilla opitaan yleensä hyvin yleispäteviä piirteitä ja mitä ylempään verkkoon siirrytään, sitä yksityiskohtaisempaa opittu tieto on. Ylempien kerrosten syötteinä käytetään siis alempien kerrosten tulosteita. Juuri verkon alimmille tasoille konvoluutiokerros usein sijoitetaan, koska nämä kerrokset käsittelevät suoraan kuvien pikseleitä. Sellaisia kuvia, joiden sisältöä ei tarvitse erikseen kertoa, eli niin sanottua luokittelematonta dataa, on saatavilla paljon ja tätä tietoa voidaan helposti käyttää alimpien kerrosten datana. Tällaisissa tilanteissa ohjaamattoman oppimisen käyttö on hyvin tehokasta. (Elements of AI, 2018e.)

Ylemmillä kerroksilla käytetään usein vastavirta-algoritmin tapaisia ohjattun koneoppimisen malleja (Elements of AI, 2018e.). Vastavirta-algoritmia eli takaisinvirtausalgoritmia käytetään kouluttamaan neuroverkkoa ja se on monen syväoppimismenetelmän ydin. Yksinkertaistettuna jokainen siirtymä verkossa eteenpäin johtaa takaisinvirtausalgoritmin säätämään mallin painoja vastaamaan haluttua tulostetta. Algoritmi sai nimensä siitä, että painoarvot päivitetään tulosteesta syötteeseen päin. (Zola & Vaughan, 2022.)

**“To allow anyone,
not to be creative
because people
are creative but to
access creativity.
And to allow the
creatives to be
even more creative
and tell even
better stories.”**

Emad Mostaque, CEO and Co-Founder
of Stability AI

GENERATIIVINEN TEKOÄLY

Generatiivinen tekoäly on laaja käsite, jota voidaan käyttää minkä tahansa tekoälyjärjestelmän kuvaamiseen, jonka ensisijainen tehtävä on luoda sisältöä kuten kuvia, tekstiä, ääntä tai koodia. Tämä on vastakohta ns. perinteisille tekoälymalleille, jotka keskittyvät tietojen analysointiin, tulkintaan ja päätöksentekoon. Niin kutsuttujen generatiivisten neuroverkkomallien saralla tapahtuneiden läpimurtojen myötä, uusia työkaluja sisällön luomiseen valokuvista ja maalauksista runoihin, koodiin, elokuvakäsikirjoituksiin ja elokuvaan on kehitelty. (Lawton, 2023.)


generatiivisia malleja ja
neuroverkoarkkitehtuureja

Tutkijat havaitsivat 2010-luvun puolivälissä uusien generatiivisten tekoälymallien potentiaalin, kun variaationaaliset autoenkooderit (VAE, variational autoencoder) generatiiviset kilpailevat neuroverkot (GAN, generative adversarial network) ja diffuusiomallit kehitettiin. Transformer-neuroverkoarkkitehtuuri esiteltiin Googlen tutkijoiden toimesta vuonna 2017. Vuonna 2020 esiteltiin NeRF-arkkitehtuuri, jonka avulla 2D-kuvista voitiin luoda 3D-sisältöä. Eri generatiiviset tekoälymallit käyttävät erilaisia tekniikoita ja lähestymistapoja uuden tiedon luomiseen. Se millainen malli tai arkkitehtuuri sopii mihinkin käyttötarkoitukseen riippuu aina käyttökohteesta. (Lawton, 2023.)

Perusmalli ja kielimalli

luonnollinen kieli
= puhuttu ja kirjoitettu kieli

Luonnollisen kielen (puhuttu ja kirjoitettu kieli) käsittely (NLP, natural language processing) ja kuvankäsittely on aiemmin tuottanut suuria haasteita tekoälyn kehityksessä, mutta syväoppimista hyödyntäen on viime aikoina saavutettu suuria edistysaskeleita juuri näillä tekoälyn sarjoilla. (Elements of AI, 2018e.) Luonnollisen kielen käsittely saa tietokoneen ymmärtämään ihmisen käyttämää kieltä. Generatiivinen tekoäly käyttää toimintansa pohjana perusmalleja (FM, foundation model), jotka ovat valtavilla tietomäärillä koulutettuja suuria koneoppimismalleja. Suuri kielimalli (LLM, large language model) on laajalla tekstiaineistolla koulutettu perusmalli, joka tulkitsee tekstiä ja hahmottaa maailmaa sen mukaan. (Tam, 2023.)

Perinteisesti tekoälyä on koulutettu hyvin tehtäväkohtaisesti, tehtäväkohtaisella datalla suorittamaan kapeaa valikoima tehtäviä. Perusmalleja koulutetaan ohjaamattomalla tavalla, valtavalla määrällä jäsentämätöntä dataa. Hienosäätämällä ne voivat mukautua monenlaisiin sovelluksiin ja loppupään tehtäviin. Toisin sanoen malli tarjoaa perustan (siitä nimitys "foundation"), jolle muita sovelluksia voidaan rakentaa. (Toner, 2023.)

Puhekielessä termiä "perusmalli" käytetään usein synonyyminä termille "suuri kielimalli", koska suuret kielimallit ovat tällä hetkellä selkein esimerkki järjestelmistä, joilla on laajoja kykyjä ja joita voidaan sovitaa tiettyihin tarkoituksiin. Oleellinen ero termeissä on se, että "suuri kielimalli" viittaa nimenomaan kieleen keskittyviin järjestelmiin, kun taas "perusmalli" pyrkii määrittelemään laajemman käsitteen. (Toner, 2023.)

Terminä suuri kielimalli on edelleen melko epämääräinen. Ei ole täysin selvää, mikä lasketaan kielimalliksi ja mikä ei: lasketaanko esimerkiksi koodilla koulutettu malli kielimalliksi? Entä sellainen malli, joka toimii pääasiassa kielen kanssa, mutta voi käyttää myös kuvia syöteinä? Ei myöskään ei ole vakiintunutta yksimielisyyttä siitä, minkä suuruinen malli lasketaan "suureksi". (Toner, 2023.)

Perusmallit toimivat sovellusten perustana. OpenAI:n suurista kielimalleista GPT-3.5:sta ja GPT-4:sta on tullut perusta monille sovelluksille lyhyen ajan kuluessa. Näistä sovelluksista OpenAI:n ChatGPT on merkittävin. ChatGPT:n ilmaisen version kielimallina toimii GPT-3.5. OpenAI siis koulutti kielimalli GPT-3.5 keskusteluilla ja edelleen ihmiskouluttajilla, luoden näin muokatun version kielimalli-GPT-3:sta, joka on erikoistunut toimimaan nimenomaan chatbot-ympäristössä. Sen jälkeen OpenAI rakensi tämän ChatGPT:hen. (Toner, 2023.)

Esimerkkejä johtavien teknologiayritysten suurista kielimalleista ovat: GPT-4 (OpenAI), BERT ja PaLM (Google) ja LLaMA (Meta AI). Näiden päälle rakennettuja keskusteluohjelmistoja, eli chatboteja ovat ChatGPT (OpenAI/GPT-3.5/4), BingChat (Microsoft/GPT-4), Google Bard (PaLM2). Esimerkiksi Googlen PaLM2:sta on olemassa lääketieteellisellä datalla hienosäädetty malli MedPaLM2 sekä kyberturvallisuusdatalla koulutettu sec-PaLM2. Tällaisia tiettyihin toimintoihin tai alaan erikoistuneita, hienosäädettyjä malleja tullaan tulevaisuudessa näkemään enenevässä määrin. StabilityAI hienosäätää malleja eri kansoille ja Googlen MedPaLM2, on testattu sairaaloissa huhtikuusta 2023 lähtien. Googlen tutkimusjohtaja Greg Corrado on avoimesti sanonut, että vaikka hän uskoo MedPaLM2:n pikkuhiljaa tulevan osaksi sellaisia terveydenhuollon tehtäviä, joissa tekoälyä voi soveltaa, hän ei haluaisi sitä osaksi omaa terveydenhuoltoaan. (Davis, 2023.)



tunnettuja kielimalleja

Perusmallien käytön edut perinteiseen tekoälyyn verrattuna ovat suorituskyvyssä ja tehokkuudessa. Ne on koulutettu niin suurella määrällä dataa, että kun niitä halutaan soveltaa perinteisesti koulutettujen tekoälymallien tavoin johonkin tiettyyn tehtäväkohtaiseen tehtävään, niitä tarvitsee enää opettaa ja hienosäätää melko pienellä määrällä tehtäväkohtaista, luokiteltua dataa. Perusmallien haitat ovat hinta: niiden kouluttaminen valtavalla tietomäärällä on kallista ja niiden käyttö- ja ylläpitokustannukset ovat suuret. Lisäksi kouluttaminen suurella datamäärällä vaikuttaa mallien luotettavuuteen. On lähes mahdotonta varmistua siitä, että käytännössä internetin eri aineistolla koulutettu malli ei sisältäisi vinoumia, vihapuhetta tai muuta haitallista tietoa. Lisäksi monien mallien koulutusmateriaalista ei edes aina ole tarkkaa tietoa. (What are Generative AI models? 2023.)

“The AI you are using is the worst and least capable AI you will ever use.”

Ethan Mollick, professor, Wharton School

Suuret kielimallit ovat mullistaneet luonnollisen kielen käsittelyn tarjoamalla malleja, jotka luovat yhtenäistä ja sujuvaa tekstiä. Luonnollisen kielen käsittelyn yhdistäminen tekoälyyn on ollut avain sille, että tekoäly on tullut koko kansan käytettäväksi. Enää ei tekoälysovellusten käyttöä varten tarvitse osata koodata tai olla tietoteknisesti lahjakas, riittää että osaa lukea ja kirjoittaa. Tämä yhdistettynä käyttäjäystävällisiin sovelluksiin selittää sen, miksi tekoäly on viimeisen vuoden ajan ollut meidän kaikkien huulilla. (Lawton, 2023.)

Suuren kielimallin rakentaminen edellyttää mallin kouluttamista valtavalla määrällä tekstiaineistoa. Kun Meta AI julkaisi ensimmäisen versionsa kielimalli LLaMa:sta helmikuussa 2023, se listasi tutkimusartikkelissaan “LLaMA: Open and Efficient Foundation Language Models” yksityiskohteisesti mallin koulutusdatan. Yli kaksi kolmasosaa mallin koulutusdatasta koostui Common Crawl -tietojoukosta, joka on vuodesta 2008 lähtien koottu ja Amazonin pilvipalveluun tallennettu, käytännössä valtava kopio internetistä. Common Crawl (CC) sisältää tekstiä monilla eri kielillä ja eri verkkotunnuksista. Englanninkielinen versio Common Crawlista (C4) on suosittu tutkimuslaboratorioiden käyttämä tietolähde. Tämän lisäksi koulutukseen oli käytetty aineistoa GitHubista (koodi), Wikipediasta, julkisesti saatavilla olevia kirjoja, ArXiv (tieteellistä dataa) ja StackExchange (korkealaatuisia kysymyksiä ja vastauksia, monipuolisesti eri aloilta tietojenkäsittelystä kemiaan). (Touvron, Lavril, Izacard, Martinet, Lachaux, Lacroix, Rozière, Goyal, Hambro, Azhar, Rodriguez, Joulin, Grave, Lample, 2023).

LLaMA =
CC: 67%
C4: 15%
GitHub: 4,5%
Wikipedia: 4,5%
Books: 4,5%
ArXiv: 2,5%
StackExchange: 2%
(Touvron & Lavril ym. 2023.)

Koulutusprosessin aikana malli oppii sanojen, lauseiden ja virkkeiden välisiä tilastollisia suhteita, mikä mahdollistaa johdonmukaisten ja kontekstuaalisesti merkityksellisten vastausten generoimisen. Tämä ei kuitenkaan tarkoita sitä, että kielimalli olisi älykäs tai että sillä olisi ymmärrystä siitä, mitä se generoi, vaan kyse on ennemminkin tilastotieteestä. (Tam, 2023.)

tekoäly on tilastotiedettä
ja todennäköisyyksiä

Julkisuuteen vuotaneiden tietojen perusteella, OpenAI:n GPT-4-mallilla on 1,76 biljoonaa parametria ollen näin yli kymmenen kertaa GPT-3:sta suurempi. Sen on myös sanottu olevan koulutettu 13 biljoonalla tokenilla (sanan osa). (Schreiner, 2023.) Parametrit ovat mallin osia, jotka se on oppinut koulutusdatasta ja jotka olennaisesti määrittelevät sen kyvykkyyden. Yksinkertaistettuna voitaisiin sanoa, että jos kyseessä olisi vieraan kielen opiskelu, olisivat parametrit vieraskielisiä sanoja, joita malli on oppinut, ja tokenit kaikki ne kirjat, joita se on aiheesta luenut. Laajan koulutusaineiston ansiosta suuret kielimallit hallitsevat useita kieliä ja erilaisia aiheita. Ne pystyvät suorittamaan erilaisia tehtäviä kuten kielenkäännös, tekstin tiivistäminen ja kysymyksiin vastaaminen, hyödyntämällä koulutusaineistosta oppittuja kielten rakenteita ja malleja. Kielimallit siis ennustavat seuraavaa sanaa tai sanan osaa ja yhdistelevät niitä toisiinsa todennäköisyyksien perusteella. (Tam, 2023.)

parametrit = opitut sanat
tokenit = kaikki luetut kirjat

Transformer-arkkitehtuuri on monien suurten kielimallien perusta. Sen ansiosta GPT:n (Generative Pre-trained Transformer) kaltaiset perusmallit pystyvät luomaan tarkemman ja asiayhteyden kannalta merkityksellisemmän tulosteen. (Lawton, 2023.)

Transformer-arkkitehtuuri

Ennen transformer-neuroverkkoarkkitehtuuria, dataa käsiteltiin sana kerrallaan siinä järjestyksessä, missä sanat esiintyvät. Sanojen välillä saattaa kuitenkin olla tärkeitäkin suhteita, vaikka ne eivät esiintyisi peräkkäin. Attention-mekanismin avulla mallit pystyvät paremmin huomioimaan nämä pitkän matkan riippuvuudet sanojen välillä. Attention-mekanismi huomioi sanojen välisiä suhteita riippumatta siitä, kuinka kaukana ne ovat toisistaan, ja määrittää, mihin sanoihin tai ilmaisuihin tekstissä on tärkeää "kiinnittää huomiota". Googlen tutkijoiden vuonna 2017 julkaistu artikkeli "Attention is all you need" esitteli täysin attention-mekanismille perustuvan transformer-arkkitehtuurin, joka on mullistanut koko tekoälyalaa viimeisen puolen vuosikymmenen ajan. (Toews, 2023.)

huomio!

Transformer-neuroverkkoarkkitehtuuri mullisti luonnollisen kielen käsittelyn, koska se pystyi aiemmista malleista poiketen ymmärtämään kokonaisia lauseita tai tekstikappaleita kerrallaan, jolloin mallin kontekstuaalinen ymmärrys edeltäjiinsä verrattuna parani huomattavasti. Transformer-mallin kyky ennustaa tekstiä perustuu kielen rakenteeseen ja redundanssiin. Transformer-malli pystyy luomaan kokonaisia tekstikappaleita generoimalla tekstiä sana kerrallaan autoregressiivisella menetelmällä eli ennustamalla tulevaa käyttäytymistä aiempien käyttäytymistietojen perusteella. Malli ei tallenna kieliopin sääntöjä, vaan se oppii ne esimerkkien avulla. Tästä syystä kielimalli pystyy tuottamaan tekstiä, joka on yhtä hyvää tai jopa parempaa kuin ihmisen tuottama teksti. (Tam, 2023.)

**“Once large
language models
are deployed,
they are black
boxes, virtually
impossible to
figure out.”**

Aleksi Kallio, manager for AI and
data analytics, CSC

Suuri kielimalli on transformer-malli suuressa mittakaavassa. Se on niin suuri, että sitä ei yleensä voi käyttää sellaisenaan yhdellä tietokoneella vaan yleensä se on palvelu, joka on saatavissa erilaisten ohjelmistojärjestelmien (API, application programming interface) tai verkkoliittymän kautta. (Tam, 2023.)

Käyttämämme chatbotit, kuten OpenAI:n ChatGPT, Googlen Bard ja Microsoftin Bing Chat, perustuvat kaikki transformer-mallille. Samoin kuvia ja videoita luovat tekoälytyökalut Midjourneysta Stable Diffusioniin ja Runwayhin, sillä niiden käyttämät diffuusio mallit hyödyntävät transformereita. Transformereiden vaikutus ulottuu kuitenkin paljon tekstiä ja kuvia syvemmälle, ja ne ovatkin avanneet uusia mahdollisuuksia monilla tieteen sarjoilla, kuten biologiassa ja robotiikassa. (Toews, 2023.)

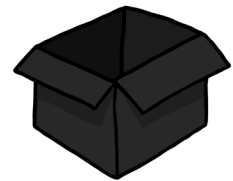
Vaikka transformerien myötä on tekoälyn saralla otettu huimia edistysaskelia, niillä on myös varjopuolensa. Transformerien aiempiin arkkitehtuureihin nähden tehokas skaalautuvuus tekee mahdolliseksi sen, että malleista voidaan rakentaa paljon aiempaa suurempia. Tällaiset suuret mallit vaativat vastaavasti massiivista laskentatehoa. (Toews, 2023.)

Transformer-mallien heikkouksiin kuuluu myös se, etteivät ne kykene oppimaan jatkuvasti. Nykyisien mallien parametrit ovat staattisia. Mallin painot (neuronien väliset yhteyksien voimakkuudet) määritetään mallin koulutuksessa, eikä näitä painoja päivitetä uuden tiedon perusteella. Transformer-pohjaiset mallien sanotaan olevan ovat "mustia laatikoita": niiden sisäinen toiminta on liian monimutkaista ja epäselvää, jotta ihminen voisi ymmärtää tarkalleen, miksi ne käyttäytyvät kuten ne tekevät. Tämä voi olla todellinen ongelma turvallisuuskriittisissä tai tiukasti säännellyissä sovelluksissa, kuten terveydenhuollossa on käytössä. (Toews, 2023.)

Transformerien käyttämä attention-mekanismi mahdollistaa sanojen välisen suhteen ymmärtämisen riippumatta siitä, kuinka kaukana sanat toisistaan ovat. Tämä tapahtuu vertaamalla jokaista yksittäistä sekvenssin sanaa jokaiseen toiseen sanaan samassa sekvenssissä. Tämän parittaisen vertailun seurauksena laskentavaiheiden määrä kasvaa nelinkertaisesti. Syötteen pituuden kaksinkertaistaminen 32 tokenista 64 tokeniin ei kaksinkertaista transformerin laskentakustannuksia, vaan nelinkertaistaa ne. Tästä syystä transformereilla saattaa joskus olla vaikeuksia käsitellä pitkiä syötteitä. (Toews, 2023.)

Suurten kielimallien sekvenssipituus eli konteksti-ikkunan koko on rajallinen. Tämä tarkoittaa sitä, että syötteen ja tulosten yhteenlaskettujen tokenien määrä ei voi ylittää tiettyä rajaa. Sekvenssipituuden kasvattaminen on aktiivinen tutkimusalue. GPT-4-mallin perusversiolle sekvenssipituus on 8192 tokenia. Muutama kuukausi sitten OpenAI julkaisi version GPT-4:stä, jonka sekvenssipituus 32 000 tokenia. OpenAI:n kilpailija Anthropic ilmoitti äskettäin uudesta mallistaan, jolla on 100 000 merkin konteksti-ikkuna. ChatGPT:n Plus-version käyttäjillä on siis käytössään 8192 tokenin konteksti-ikkuna ja ilmaisversion käyttäjillä 4096 tokenin konteksti-ikkuna. Suuremmat konteksti-ikkunat OpenAI:lta saa käyttöönsä joko API- tai Enterprise-käyttäjänä. (Toews, 2023.) Mallien hinta käyttäjälle määritellään usein juuri tokenien perusteella.

transformer-malli:



2048 tokenia ~ 1500 sanaa (engl.)
4096 tokenia ~ 3000 sanaa (engl.)
8192 tokenia ~ 6000 sanaa (engl.)

Transformer on poikkeuksellinen tekoälyarkkitehtuuri ja ne ovat muodostuneet modernin tekoälyn perustaksi. Niiden rajoitukset huomioon ottaen on kuitenkin selvää, että entistä edistyksellisempiä tekoälyarkkitehtuureja kehitellään. (Toews, 2023.)

Miten kaikki tapahtuu?

ihminen → algoritmi → malli

Yksinkertaistettuna malli on tietokoneohjelma. Se on joukko ohjeita, jotka suorittavat erilaisia laskutoimituksia syöteaineistolla ja antavat tulosteen. Koneoppimis- tai tekoälymallin erityisyys piilee siinä, että ihmisohjelmoijat eivät kirjoita ohjeita suoraan, vaan sen sijaan he kirjoittavat joukon ohjeita eli algoritmin, joka tarkastelee suuria määriä olemassa olevaa dataa ja määrittää mallin sen pohjalta. Näin ollen ihmisohjelmoijat eivät rakenna mallia, vaan rakentavat algoritmin, jonka perusteella malli rakennetaan. Ohjelmoijat määrittelevät siis mallin arkkitehtuurin ja säännöt. He eivät kuitenkaan luo neuroneita tai neuroneiden välisiä painoja. Tämä tehdään prosessissa, jota kutsutaan kouluttamiseksi. Kouluttamisessa malli noudattaa algoritmin ohjeita ja määrittelee muuttujat itse. (Muehmel, 2023.)

Suurista kielimalleista puhuttaessa tarkasteltava tieto on tekstiä. Koulutusprosessin aikana malli tarkastelee tekstiä ja yrittää tuottaa omaa tekstiä tulosteena. Aluksi tuloste on sekavaa, mutta kokeilun ja erehtymisen sekä jatkuvan tulosteen ja syötteen välisen vertailun avulla tulosteen laatu paranee vähitellen ja teksti muuttuu ymmärrettävämmäksi. (Muehmel, 2023.)

Riittävän ajan, tietokoneressurssien ja koulutusaineiston avulla malli ”oppii” tuottamaan tekstiä, joka ei eroa ihmisen tuottamasta. Vahvistusoppimisen avulla, jossa ihmiset antavat mallille palautetta sen tekstistä (RLHF, reinforcement learning from human feedback) malli parantaa itseään saamansa palautteen perusteella. (Muehmel, 2023.)

Koulutusprosessin tuloksena syntyy neuroverkko, jossa on satoja miljardeja yhteyksiä, eli painoja miljoonien neuronien välillä. Jokainen paino ja jokainen neuroni on mallin itse määrittelemä matemaattinen kaava. Kielimallit toimivat tokeneilla (sanan osilla eli merkkiyksiköillä) ja vaativat merkkiyksikkömuuntimen muuttamaan tekstin tokeneiksi, joita malli voi käsitellä. Tokenit ovat kokonaislukuja, jotka vastaavat sanoja tai sanan osia. Englannin kielessä yksi token vastaa yleensä yhtä sanaa tai välimerkkiä ja 100 tokenia on noin 75 englannin kielistä sanaa. Yksi token vastaa

1 token ~ 4 merkkiä (engl.)
1 token ~ 3/4 sanaa (engl.)
100 tokenia ~ 75 sanaa (engl.)



pääsääntöisesti noin 1–4 kirjainta. Jos merkkiyksikkömuunnin on luotu englannin kielelle, se toimii luonnollisesti huomattavasti paremmin muilla kielillä. Kuten vieressä olevasta esimerkistä huomaa, Open AI:n englannin kielelle luotu tokenizer toimii hyvin englanniksi, mutta suomen kielellä se ei ole kovin tehokas. Oman mallin rakentaessaan, voi muokata merkkiyksikkömuuntimen haluamallaan tavalla. Tokenien ei tarvitse edes vastata luonnollisen kielen osia, ne voivat olla esimerkiksi sointuja, mikäli malli tuottaa musiikkia. (Muehmel, 2023; Lukianov 2023, Open AI, 2023.)

Käyttäjä on vuorovaikutuksessa mallin kanssa käyttäjäliittymän välityksellä. Esimerkiksi ChatGPT on käyttäjäliittymä, jonka välityksellä käyttäjä on vuorovaikutuksessa GPT-3.5 tai 4-mallin kanssa. GPT-3.5 ja GPT-4-mallit ovat kielimalleja, jotka on optimoitu keskustelurajapintoja varten. Käyttäjä antaa mallille kehoitteen tekstinä, jonka käyttäjäliittymän taustalla olevat malli muuntaa tokeneiksi. (Muehmel, 2023.)

Tämän jälkeen malli ryhtyy luomaan vastausta, joka kuulostaa oikealta perustuen siihen valtavaan tekstimäärään, jonka sille on koulutusprosessissa näytetty. On tärkeää huomata, että mallilla ei ole muistia, vaan se alkaa generoida vastausta token kerrallaan. Jokaisen generoimansa tokenin jälkeen se suorittaa laskennan uudelleen seuraavaa generoimansa tokenia varten. Token tokenilta malli tarjoaa annettuun kehoitteen perusteella joukon mahdollisia seuraavia tokeneja. Malli valitsee näistä vaihtoehdoista vastaukseensa yleensä tokenin, jolla on korkea, mutta ei välttämättä korkein todennäköisyys kuulostaa oikealta. Erilaisia keinoja käyttäen mallia voidaan ohjata satunnaisuuteen ja luovuuteen eli valitsemaan vähemmän todennäköisen tokenin. (Muehmel, 2023; Kleczek, 2023.)

Tämä
Tämä on
Tämä on kirja
Tämä on kirja, joka
Tämä on kirja, joka kertoo
Tämä on kirja, joka kertoo tarinan.

Keskusteluhistoriaan lisätään jokainen keskusteluikkunan kehote ja vastaus. Parempien vastausten generoimiseksi jokaisen uuden kehoitteen myötä, keskusteluikkunan koko keskusteluhistoria lähetetään mallille tarkasteltavaksi. Ajan mittaan keskusteluhistorian tokenien määrä voi ylittää mallin sekvenssipituuden. Kun sekvenssipituus ylitetään, malli saattaa alkaa hallusinoimaan, koska se ei enää voi käyttää koko keskusteluhistoriaa aineistonaan. (Dunn, 2023.) Tästä syystä on hyvä tietää käyttämänsä mallin sekvenssipituus, tiedostaa mikä token on ja ymmärtää, kuinka tokenien määrä keskusteluhistoriassa vaikuttaa mallin käyttöön ja toimivuuteen.

“All models are wrong, some are useful.”

George E.P. Box, statistician

ChatGPT tai mikään muukaan kielimalleihin perustuva chatbot ei voi taata, että sen tuottama tulos on faktuaalisesti oikea, vaan ainoastaan sen, että se kuulostaa oikealta. Sen vastaukset eivät perustu muistissa oleviin tietoihin, vaan ne luodaan lennosta perustuen siihen, miten mallin painot on määritelty ja mitä kehoitteessa ja keskusteluikkunassa on käsitelty. Kielimallien vahvuus ei ole tiedon muistaminen tai hakeminen vaan tekstien luominen. (Muehmel, 2023.)

Suurien kielimallien tietämys rajoittuu niiden koulutusaineistossa oleviin tietoihin. Esimerkiksi ChatGPT:n ilmaisversion koulutusaineisto päättyi syyskuuhun 2021. Lisäksi tiedetään, että monien kielimallien koulutusaineisto on suuressa määrin julkisesti saatavilla olevaa tietoa internetistä. Koulutusaineisto voi siis olla puutteellista, virheellistä, vinoutunutta tai vanhentunutta. Tästä huolimatta chatbotit pyrkivät aina antamaan vastauksen, joka kuulostaa oikealta. (Muehmel, 2023.)

Yksi tapa lieventää tätä puutetta on tarjota kielimallille kehoitteessa yksityiskohtaisen kysymyksen lisäksi lisätietoa aiheesta. Kielimalleja voidaan myös laajentaa lisäosilla (plugins), jotka antavat niille pääsyn ulkoisiin tietolähteisiin ja internetiin, laajentaen niiden päättelykykyä ja mahdollistaen niiden toiminnan todellisessa maailmassa. (Muehmel, 2023.)

Tekoälytaide

Tekoälytaide mahdollistui, kun ensin internetin myötä opetusaineistoksi tarjoutui lähes rajaton määrä kuvia ja lopuksi luonnollista kieltä ymmärtävät kielimallit valjastettiin neuroverkkoihin (Tam, 2023.). Aidonnäköisten kuvien luomiseksi kehitettiin generatiivinen kilpaileva neuroverkko (GAN), jonka esiteltiin ensikertaa vuonna 2014. Tässä kaksi neuroverkkoa asetetaan kilpailemaan keskenään. Toinen verkko opetetaan generoimaan opetusdataa muistuttavia kuvia, kun toinen verkko taas opetetaan

tunnistamaan generoidut kuvat opetusdatan kuvista. GAN-verkon muodostamaa kahta verkkoa opetetaan yhtäaikaisesti ja ne oppivat molemmat vähitellen paremmiksi omassa tehtävässään. Vähitellen generoidut kuvat eivät ole vain toisintoja opetusdatan kuvista, vaan verkko oppii tuottamaan kokonaan uusia, aidonnäköisiä kuvia. Vaikka GAN:t tuottavat uusia kuvia, niiden lopputuotokset heijastelevat kuitenkin niille syötettyä dataa. (Elements of AI, 2018e.)

Diffuusiomallit toimivat tuhoamalla koulutusdataa lisäämällä siihen melua. Malli oppii palauttamaan datan tekemällä saman prosessin käänteisesti. Toisin sanoen, diffuusiomallit generoivat kuvia melusta. Esimerkiksi kuvan-generointipalvelut Dall-E 2, Stable Diffusion ja Midjourney perustuvat diffuusiomalleille, jotka on yhdistetty suuriin kielimalleihin. Diffuusiomalleja käytetään myös äänen, videon ja 3D-sisällön tuottamiseen. Lisäksi diffuusiotekniikkaa voidaan käyttää datan imputointiin, missä puuttuva tieto ennustetaan ja generoidaan datajoukon sisäisten tilastollisten suhteiden perusteella. (Lawton, 2023.) Diffuusiomallit esiteltiin ensikertaa vuonna 2015 ja ne oli laajasti saatavilla vuoden 2022 loppuun mennessä. Diffuusiomallit ohittivat GAN:t kustannustehokkuudessa ja hienostuneisuudessa ja pystyivät luomaan uusia kuvia yksinkertaisista tekstikehoteista. (Grady & Castro, 2023.)

On tärkeää huomata, että mallit ja arkkitehtuurit ovat jatkuvan kehityksen alla. Tutkijat pyrkivät parantamaan yksittäisiä malleja sekä tapoja yhdistää niitä muihin malleihin ja prosessointitekniikoihin. On luultavaa, että tulevaisuudessa entistä tehokkaammat mallit tulevat syrjäyttämään nykyiset mallit. (Lawton, 2023.) On myös tärkeää muistaa valita kuhunkin käyttötarkoitukseen sopiva työkalu. ChatGPT:tä ei tule käyttää Googlen sijaan, koska ChatGPT ei ole hakukone, joka etsii Internetistä sivuja.

ChatGPT ≈ Google-haku

“AI is just an extension of human history as a whole.”

Tuan Ho, CEO &
Cofounder of The Point AI

Teknologiajättien kilpajuoksu

OpenAI aloitti toimintansa vuonna 2015 voittoa tavoittelemattomana yrityksenä. Sen tavoitteena oli rakentaa avoimen lähdekoodin tekoälyohjelmistoja läpinäkyvästi, jotta koko maailma voisi niistä hyötyä. OpenAI käytti poikkeuksellisen suuria tietomääriä ja tehokkaita ohjelmistoja tekoälytuotteiden luomiseen. Suuret kustannukset vaikeuttivat yrityksen johtamista voittoa tavoittelemattomana ja vuonna 2019 OpenAI muutettiin voittoa tavoittelevaksi yritykseksi. (Verma, 2023.)

OpenAI nousi valtavaan suosioon julkaistuaan suuren kielimallinsa suuren yleisön käytettäväksi. Yhtiö on kuitenkin saanut paljon palautetta siitä, että nimestään ja taustastaan poiketen sen toiminta on muuttunut jatkuvasti vähemmän läpinäkyväksi ja avoimeksi. On kuitenkin kiistatta selvää, että OpenAI on ottanut viime vuosina vaikuttavia edistysaskelia tekoälyn ja kielimallien saralla skaalamaalla olemassa olevia koneoppimisalgoritmeja mittasuhteisiin, joista aiemmin ei osattu edes uneksia. (Knight, 2023.)

Maaliskuussa 2022 Google DeepMindin (Alphabet Inc:n tytäryhtiö ja tekoälytutkimuslaboratorio) tutkijat julkaisivat ”Training Compute-Optimal Large Language Models”-tutkimusraportin. Sen mukaan mallin ja koulutusaineiston tulisi skaalautua samassa suhteessa – mallin kaksinkertaistuksessa siis myös koulutusaineiston tulisi kaksinkertaistua. Tämän hypoteesin testaakseen tutkijat kouluttivat laskentaoptimaalisen mallin Chinchillan, joka käytti samaa laskentabudjettia (compute budget) kuin Googlen 280

miljardin parametrin kielimalli Gopher. Chinchillassa on Gopherista poiketen vain 70 miljardia parametria, mutta neljä kertaa enemmän koulutusdataa. Chinchilla osoitti suoriutuvansa monista tehtävistä vähintäänkin yhtä hyvin kuin esimerkiksi Gopher tai GPT-3. Chinchilla siis käyttää huomattavasti vähemmän laskentaa hienosäätöön ja päättelyyn, mikä helpottaa merkittävästi sen jälkikäyttöä: sen pienempi koko tekee päätte-lystä halvempaa ja hienosäätämisestä helpompaa. (Hoffmann, Borgeaud, Mensch, Buchatskaya, Cai, Rutherford, de Las Casas, Hendricks, Welbl, Clark, Hennigan, Noland, Millican, van den Driessche, Damoc, Guy, Osindero, Simonyan, Elsen, Rae, Vinyals & Sifre. 2022.)

parametrit = opitut sanat
tokenit = kaikki luetut kirjat

Huhtikuussa 2023 OpenAI:n Altman totesi, että tulevaisuudessa tekoäly-ala ei enää edisty malleja suurentamalla, vaan niitä parannellaan toisilla tavoin. Cohere-yhtiön perustaja Nick Frosst yhtyi Altmanin ajatuksiin ja totesi, että uudet tekoälyarkkitehtuurit sekä edistyneempi, ihmispalauttee-seen perustuva hienosäätö ovat lupaavia tutkimussuuntia alalla. (Knight, 2023.)

ChatGPT laukaisi tekoälykilpailun suurten teknologiajättien, kuten Microsoftin, Googlen ja OpenAI:n välillä. Kilpailun seurauksena avoi-muus suurten kielimallien osalta väheni, kun esimerkiksi lähdekoodista, malliarkkitehtuurista, painoista ja koulutusdatasta tuli yksityislaishuuksia. OpenAI ei ole paljastanut julkisuuteen juuri tietoa GPT-4:n arkkitehtuu-rista. Sam Altman on sanonut prosessin maksaneen ”yli 100 miljoonaa dollaria”. The tiny corp -yhtiön perustaja ja hakkeri George Hotz paljasti kesäkuussa 2023, että GPT-4 on yhden mallin sijaan luultavasti yhdis-telmä kahdeksaa eri mallia, joista jokaisessa on 220 miljardia parametria. Jos Hotzin väite yhdistelmämallista pitää paikkansa, tarkoittaisi se teknolo-gian todella olevan siirtymässä mallien suurentamisesta toiseen suuntaan. (Knight, 2023; Commoditizing the Petaflop – with George Hotz of the tiny corp, 2023.)

Samaan aikaan kun suurten teknologiayrityksien isännöimät chatbotit ja tekoälyavustajat kuten ChatGPT, Bard, Bing ja Copilot tekevät aiemmin testilaboratorioissa rajoitetusti saatavilla olleesta teknologiasta helpom-min saavutettavan, avoimen lähdekoodin yhteisö on vastannut kilpailuun omilla kielimalleillaan. Viime kuukausina useita avoimen lähdekoodin kielimalleja on julkaistu, jotka haastavat koko Big Technin perustaman liike-toimintamallin. (Dickson, 2023.) Big Tech, joka tunnetaan myös nimellä Tech Giants tai Big Five viittaa ryhmään teknologiayrityksiä, jotka ovat hallinneet alaa vuosia kokonsa, vaikutusvaltansa ja taloudellisen menes-tyksensä vuoksi. Termillä tarkoitetaan tyypillisesti maailman viittä suurinta teknologiayritystä: Alphabet (Google), Amazon, Apple, Meta (Facebook) ja Microsoft. (Shinde, 2023.)

Big Tech/Tech Giants/Big Five

Avoimen lähdekoodin kielimallivaihtoehdot ovat osoittautuneet olevan kilpailukykyisiä erityisesti kolmella tasolla. Muutaman miljardin parametrin suu-ruiset kielimallit voivat kilpailla erittäin suurten mallien kanssa suori-tuskyvyyden suhteen, mikäli ne ovat koulutettu suurilla tietojoukoilla. Pienempiä malleja voi hienosäätää erittäin pienellä budjetilla ja vaatimattomalla tietomäärällä. Lisäksi avoimen lähdekoodin suurten kielimallien kehitys on nopeampaa kuin suljetussa ekosysteemissä. (Dickson, 2023.)

“We need to keep CEOs away from AI regulation.”

Marietje Schaake,
international policy director at
Stanford University’s Cyber Policy Center

Meta kehitti aiemmin tänä vuonna suuren kielimallinsa LLaMa:n, joka pystyy toimimaan chatbotin taustalla, aivan kuten OpenAI:n GPT-mallit sekä Googlen PaLM. Toisin kuin OpenAI tai Google, Meta teki mallinsa lähdekoodin avoimeksi, joten sitä voidaan kopioida ja muokata uusien chatbottien luomiseksi. Metan mukaan teknologian lähdekoodin avoimuus mahdollistaa kaikille tutkijoille tasavertaisen mahdollisuuden tekoälyn tutkimiseen. Kriitikot pelkäävät, että avoin lähdekoodi helpottaa tekniikan manipulointia sekä disinformaation ja vihapuheen levittämistä. (Jones, 2023.)

Tällaisia avoimen lähdekoodin suuria kielimalleja ovat esimerkiksi edellä mainittu Metan ei-kaupallisiin tarkoituksiin kehittämä suurien kielimallien perhe LLaMa sekä yliopistojen tutkijoiden LLaMa:n päälle luomat mallit Alpaca ja Vicuna, jotka toimivat kuten ChatGPT, tai StabilityAI:n StableLM. Technology Innovation Instituten (TII) kehittämää avoimen lähdekoodin kielimalli Falconia voi käyttää kaupallisiin tarkoituksiin, samoin kuin Open Assistant-mallia, joka on avoin niin koodin, mallin painojen kuin koulutusdatankin suhteen. (Dickson, 2023.) Todellisuudessa avoimen lähdekoodin malli ei tarkoita kuitenkaan aina sitä, että kaikki mallin tiedot olisivat julkisia vaan avoimuus on melko vaihtelevaa. Siksi niitä olisikin ehkä tarkoituksenmukaisempaa kutsua avoimen lisenssin malleiksi.

lisenssin
avoimen lähdekoodin malli

Metan LLaMa-mallin kaupallinen versio LLaMa2 julkaistiin heinäkuussa 2023. Se on koulutettu kahdella biljoonalla tokenilla ja sen mallien koko vaihtelee 7–70 miljardin parametrin välillä (Meta AI, 2023). Metan kansainvälisten asioiden päällikkö Nick Cleggin mukaan avoimuus on paras ratkaisu tekoälyyn liittyvien pelkojen lievittämiseen. Cleggin mukaan avoimuutta voidaan edistää yhteistyöllä ja tekniikan stressitestauksella

(Dickson, 2023.). Avoimen lähdekoodin malleista puhuttaessa ei tule unohtaa TurkuNLP-tutkijaryhmä helmikuussa julkaisemaa, täysin suomenkielistä FinGPT3-kielimallia. (TurkuNLP, 2023.).



Mitä tämä tarkoittaa suurille teknologiayrityksille? Googlelta julkisuuteen keväällä 2023 vuotaman dokumentin perusteella teknologian kilpailuedun säilyttäminen on entistä vaikeampaa nyt, kun suurien kielimallien huippututkimus on edullisempaa. Dokumentin kirjoittajan mukaan Google voi yrittää pitää tiukasti kiinni salaisuuksistaan tai pysyä mukana kilpailussa oppimalla muilta ja jakamalla tietoaan. (Dickson, 2023; Patel & Ahmad, 2023b.)

On tärkeää ymmärtää kuinka yritykset tienaavat näillä työkaluilla. Esimerkiksi OpenAI saa rahaa API-lisensseistä, ChatGPT Plus-tilauksista, sijoituksista ja yhteistyöstä eri toimijoiden kanssa. OpenAI myöntää lisenssejä yrityksille ja organisaatioille, jotka käyttävät kielimallia virtuaalivestustajien, chatbottien ja muiden sovellusten luomiseen. OpenAI saa myös satunnaisesti rahoitusta yksityisiltä sijoittajilta ja yrityksiltä vastineeksi omistussuosuksista. Esimerkiksi Microsoft on sijoittanut yritykseen miljardeja dollareita. Lisäksi organisaatio voi saada rahoitusta myös muilta organisaatioilta ja tahoilta, joiden tarkoituksena on pääasiassa tukea OpenAI:n missiota. OpenAI on tehnyt myös yhteistyötä yritysten kanssa auttaakseen näitä tunnistamaan tekoälyyn liittyviä haasteita ja löytääkseen mahdollisia ratkaisuja niihin. Näissä tapauksissa OpenAI saa korvausta tarjoamistaan palveluista ja toisinaan myös osuuden tekoälyteknologian yritykselle tuottamasta voitoista. On kuitenkin hyvä muistaa, että tekoälytyökalujen kuten ChatGPT:n ylläpito ja käyttö, mallien koulutuksesta puhumattakaan, ovat varsin kalliita.

Kuka tienaa?
Miten tienaa?

Pilvipalveluiden tarjoajat voivat saavuttaa potentiaalisen kilpailuedun tarjoamalla hyvää suurta kielimallia tekoälysovellusten kehittämiin. Esimerkiksi Microsoft ilmoitti heinäkuussa 2023 tekevänsä Meta:n LLaMa2:n saataville Azure-pilvipalvelussaan. Meta:n edustajan mukaan yritys ei keskity rahastamaan LLaMa2:lla, vaan yritykset, jotka haluavat käyttää ohjelmistoa kaupallisen lisenssin kautta, maksavat Microsoftille ylläpidosta. Vaikka Meta ei suoraan hyödy rahallisesti sopimuksesta Microsoftin kanssa, sillä voi olla potentiaalista hyötyä saadessaan tekoälyohjelmistoaan laajemmin käyttäjien ja kehittäjien käsiin. LLaMa2 tulee olemaan saatavilla myös Amazon Web Servicesin ja Hugging Face-palveluiden kautta. Hugging Face on suosittu palvelu tekoälytutkijoiden keskuudessa. Sopimus on merkittävä Microsoftille, koska se osoittaa, että se on kiinnostunut myös muista, kuin OpenAI:n kehittämistä kielimalleista. (Vanian, 2023.)

Kilpailu generatiivisen tekoälyn saralla on kovaa. Kiinalainen verkkokauppajätti Alibaba ilmoitti elokuussa 2023 tuovansa huhtikuussa lanseeraamansa suuren kielimallinsa Tongyi Qianwen kolmansien osapuolien saataville, pyrkien näin lisäämään tuotteensa käyttöä ja ottamaan johtavaa roolia teknologiassa. Tongyi Qianwen on saatavilla eri kokoisina malleina seitsemästä miljardista parametrilla ylöspäin ja se mahdollistaa sisällön generoinnin englanniksi ja kiinaksi. Alibaban avoimen lähdekoodien mallit tulevat olemaan Qwen-7B sekä keskustelusovelluksille suunnitellun versio Qwen-7B-Chat. Jakelu auttaa yhtiötä saamaan lisää käyttäjiä tekoälymallilleen. Kiinan suurin verkkokauppayhtiö pyrkii vahvistamaan pilvipalveluosastoaan investoimalla tekoälyyn ja tunnistamalla pilvipalvelut kriittisenä tulevaisuuden kasvualueena. Tongyi Qianwen päälle on jo rakennettu sovelluksia, esimerkiksi heinäkuussa kiinalainen yritys julkaisi tekoälypalvelu Tongyi Wanxiangin, joka voi generoida kuvia kiinan- ja englannin kielisistä tekstisyötteistä. Myös kiinalainen teknologiayritys Baidu on julkaissut oman perusmallinsa Baidu Earnien ja maaliskuussa se julkaisi Baidu Earnien päälle rakennetun Earnie Botin. (Kharpal, 2023a & 2023b.)

StabilityAI:n Emad Mostaquen mukaan on luultavaa, että Google, Microsoft, OpenAI ja StabilityAI ovat yritykset, jotka tulevaisuudessa rakentavat perusmalleja muiden yritysten käyttöön. Pilvipohjaisten kielimallien markkinat eivät tule katoamaan, koska perusmallien esikoulutus on edelleen kallista eikä avoimen lähdekoodin suuret kielimallit sovi kaikille. Microsoftin ja Googlen kaltaisilla yrityksillä on myös erittäin vahvat jakelukanavat sovellusten ja asiakaskuntansa kautta. Avoimen lähdekoodin yhteisön ponnistelut laajentavat markkinoita mahdollistaen esimerkiksi suurien kielimallien käytön uusissa ympäristöissä, kuten ilman palvelinta omassa kannettavassa tietokoneessa. Esimerkiksi GPT4ALL on Nomic AI:n luotsaama projekti, joka tähtää siihen, että voisimme käyttää suuria kielimalleja omilla laitteillamme (Nomic, 2023.). Avoimen lähdekoodin suuret kielimallit myös jossain määrin kaupallistavat markkinoita ja pakottavat teknologiajättiäiset tarjoamaan asiakkailleen kilpailukykyisempiä hintoja. (Dickson, 2023; Why AI Matters And How To Deal With The Coming Change w/ Emad Mostaque, 2023).

**“AI comes with
both opportunities
and threats
and the balance
will depend
on our decisions.”**

European Commission

Tekoälyteknologioiden nopean kehityksen vuoksi tekoälyn sääntely on viime vuosina noussut keskeiseksi poliittiseksi kysymykseksi koko maailmassa. Vuonna 2020 Euroopan komissio sitoutui edistämään tekoälyn käyttöönottoa ja kartoittamaan teknologian käyttöön liittyviä riskejä. Vaikka Euroopan komissio alun perin lähestyi tekoälykysymystä ei-sitovan, pehmeän lainsäädännön ja suositusten kautta, on se sittemmin siirtynyt lainsäädännölliseen lähestymistapaan ja vaatinut yhtenäisiä sääntöjä tekoälyjärjestelmien kehittämiseksi, markkinoille saattamiselle ja käytölle. (European parliament, 2023.) Lakien hankaluus on siinä, ettei niitä pystytä säätämään samaa vauhtia kuin teknologia kehittyy. Eri valtioilla on erilaisia näkö- ja tulokulmia tekoälyhaasteisiin. Euroopan unioni (EU) esimerkiksi lähestyy tekoälyä riskien näkökulmasta.

EU on laatinut tekoälyä koskevan toimenpidepaketin, joka sisältää tekoälyasetuksen, vastuukysymyksiä koskevan direktiiviehdotuksen ja alakohtaisen tuoteturvallisuuksääntöjen tarkistamisen. Toimenpidepaketin avulla EU yrittää minimoida turvallisuusriskit ja suojella perusoikeuksia, sekä valmistautuu tekoälyn käytöstä mahdollisesti aiheutuvien vastuukysymysten ratkaisemiseen. Euroopan parlamentti hyväksyi kesäkuussa 2023 kantansa tekoälyasetuksesta. Tämän myötä lain lopullista muotoa koskevat neuvottelut voitiin aloittaa. Odotettavissa on tekoälyjärjestelmien määrittelyn tarkistaminen ja kiellettyjen tekoälyjärjestelmien luettelon laajentaminen. Valmisteilla olevien lakien ja asetusten lisäksi tekoälyyn pohjautuvien järjestelmien valmistamista ja käyttöä koskevat jo monet olemassa olevat säädökset, joskin niiden soveltuvuuteen liittyy paljon epävarmuustekijöitä. (Euroopan parlamentti, 2023; Merilampi 2023a & Merilampi 2023b).

Asetusehdotuksessa tekoälyä hyödyntävät järjestelmät on jaoteltu neljään riskitasoon: kielletyt ja korkean riskin sekä rajoitetun ja vähäisen riskin tekoälyjärjestelmät. EU:n asetusehdotuksessa kiellettyihin järjestelmiin kuuluvat sellaiset järjestelmät, joiden käyttö loukkaa perusoikeuksia ja on täten EU:n arvojen vastaisia. Ehdotuksen mukaan tällaiset järjestelmät aiheuttavat merkittävän riskin terveydelle, turvallisuudelle ja perusoikeuksille. Näitä järjestelmiä voisivat olla esimerkiksi ihmisten valvonta ja

kielletyt
korkean riskin
vähäisen riskin
minimaalisen riskin

} AI-järjestelmät

tunnistaminen julkisilla paikoilla tai pisteytys sosiaalisen toiminnan perusteella sekä ennustava poliisitoiminta. (Euroopan parlamentti, 2023; Merilampi 2023b.)

Korkean riskin järjestelmiin kuuluisivat esimerkiksi lainan myöntämisessä, opiskelijavalinnoissa ja henkilöstöhallinnassa käytettävät sekä vaaleissa äänestäjiin vaikuttavat tekoälyjärjestelmät, joilla voi olla kielteinen vaikutus ihmisten turvallisuuteen tai perusoikeuksiin. Korkean riskin tekoälyjärjestelmien tarjoajien olisi rekisteröitävä järjestelmänsä komission hallinnoimaan EU:n laajuiseen tietokantaan ennen kuin niitä saisi tuoda markkinoille tai ottaa käyttöön. (Euroopan parlamentti, 2023; Merilampi 2023b; European Parliament, 2023.)

Sääntely, joka koskee rajoitetun tai vähäisen riskintason tekoälyjärjestelmiä on lähinnä läpinäkyvyysvaatimuksia, joiden on tarkoitus taata, että käyttäjä on tietoinen olevansa tekemisissä tekoälyn kanssa. Sisällön ollessa tekoälyn tuottamaa, tulee asian tulla selkeästi ilmi. Tähän kategoriaan kuuluisivat esimerkiksi ihmisten kanssa vuorovaikutuksessa olevat järjestelmät kuten chatbotit, tunteiden tunnistamiseen kehitetyt järjestelmät, biometriset kategorisointijärjestelmät sekä tekoälyjärjestelmät, jotka tuottavat tai manipuloivat kuvia, ääntä tai videomateriaalia. (Euroopan parlamentti, 2023; Merilampi 2023b; European Parliament, 2023)

Tekoälyn sääntelytavat ovat maailmanlaajuisesti melko kirjavat. Yhdysvallat suhtautui tekoälyyn aluksi melko väljästi, mutta tarve sääntelyyn on viime aikoina kasvanut. Kiinan Cyberspace-hallinto on neuvotellut ehdotuksesta tekoälyn sääntelyyn, kun taas Iso-Britannia lähestyy asiaa innovaatioystävällisten sääntelyperiaatteiden näkökulmasta. Kansainvälisellä tasolla Taloudellisen yhteistyön ja kehityksen järjestö (OECD) hyväksyi tekoälysuosituksen ei-sitovana vuonna 2019, UNESCO antoi eettisiä suosituksia tekoälystä vuonna 2021, ja Euroopan neuvosto työskentelee parhaillaan kansainvälisen tekoälyä koskevan yleissopimuksen parissa. Lisäksi EU:n ja Yhdysvaltojen välisessä uudessa teknologiakumppanuudessa (Trade and Technology Council) EU ja USA pyrkivät löytämään yhteisen sävelen luotettavan ja vastuullisen tekoälyn periaatteista. (European parliament, 2023.)

“The future of AI is not as grim as some people think or as rosy as others think. The risks are real, but I am optimistic that they can be managed.”

Bill Gates, computer programmer and entrepreneur

EU:n ehdotus edellyttää jäsenvaltioita nimeämään toimivaltaisia viranomaisia, kuten kansallisen valvontaviranomaisen, joiden tehtävänä on valvoa asetuksen soveltamista ja täytäntöönpanoa, sekä luomaan EU:n tasolla toimivan Euroopan tekoälylautakunnan. Kansalliset markkinalvontaviranomaiset olisivat ehdotuksen mukaan vastuussa korkean riskin tekoälyjärjestelmätoimijoiden lain noudattamisen arvioinnista ja korjaavien toimenpiteiden tekemisestä. Heillä olisi pääsy luottamuksellisiin tietoihin, kuten esimerkiksi tekoälyjärjestelmien lähdekoodeihin. Jäsenvaltioiden olisi säädettävä rangaistuksia koskevat säännöt ja toteutettava kaikki tarvittavat toimenpiteet varmistaakseen niiden asianmukaisen ja tehokkaan täytäntöönpanon. (European parliament, 2023.)

Komissio ehdottaa myös, että jäsenvaltiot tai Euroopan tietosuojavaltuutettu voisi perustaa hallitun ympäristön, niin sanotun hiekkalaatikon, joka helpottaisi innovatiivisten tekoälyjärjestelmien kehittämistä, testaamista ja validointia ennen markkinoille saattamista. Tähän osallistujilla olisi esimerkiksi mahdollisuus käyttää henkilötietoja tekoälyn innovaation edistämiseksi GDPR:n vaatimusten mukaisesti. Muita toimenpiteitä on suunnattu erityisesti pienimuotoisille palveluntarjoajille ja start-up-yrityksille. (European parliament, 2023.)

EU:n tekoälyasetus on herättänyt keskustelua eri sidosryhmien kesken. Big Data Value Association, teollisuuden vetämä kansainvälinen voittoa tavoittelematon järjestö, kritisoi muun muassa asetuksen määritelmää tekoälyjärjestelmille. Sen mukaan määritelmä on varsin laaja ja kattaa paljon enemmän kuin mitä subjektiivisesti ymmärretään tekoälyksi. Täten myös yksinkertaisimmat hakualgoritmit, lajittelualgoritmit ja reititys-algoritmit kuuluisivat EU:n tekoälyn määritelmän piiriin, ja olisi siten uusien sääntöjen alaisia. Lisäksi esimerkiksi AmCham, American Chamber of Commerce in the EU, ehdottaa ylisääntelyn välttämistä hyväksymällä kapeamman tekoälyjärjestelmien määritelmän. Täten keskityttäisiin pääasiassa korkean riskin tekoälysovelluksiin laajentamatta sitä muihin, pienen riskin tekoälysovelluksiin tai -ohjelmistoihin. (European parliament, 2023.) Suomen tekoälykeskus Finnish Center for Artificial Intelligence (FCAI) on myös osoittanut huolensa säätelystä. FCAI:n mukaan, säätelyn ollessa huonosti toteutettu tai väärin mitoitettu se voi merkittävästi hidastaa teknologista ja taloudellista kehitystä Euroopassa. (FCAI, 2023). Painopisteen ei pitäisi olla keskittämisessä ja kontrolloinnissa, vaan ulottuvuudessa ja yhteistyössä.

keskittäminen ja kontrolli
vs.
ulottuvuus ja yhteistyö

Tekoälyn kehitys on muuttanut yhteiskuntaa vallankumouksellisesti. Sillä on potentiaalia tehostaa työskentelyä ja vapauttaa aikaa, mutta sillä on myös varjopuolensa. Tekoäly tulee integroitumaan yhä useampaan arkipäiväiseen sovellukseen ja siksi on hyvä tiedostaa sen olemassaolo, mahdollisuudet, heikkoudet sekä sen vaikutus elämäämme. (Hursti ym. 2023a.) Mallit voivat tuottaa erilaisia vastauksia samoille kysymyksille, mikä vaikeuttaa käyttäjän kykyä arvioida tulosten tarkkuutta ja luotettavuutta – generatiivinen tekoäly haastaa kykymme selittää, miten vastaukset tuotetaan (Chui, Roberts ym., 2023.).

minimoidaan riskit

Tekoälyn etiikassa tarkastelun kohteena on se, miten eri toimijoiden tulisi käyttäytyä, jotta minimoitaisiin tekoälyn suunnittelusta, soveltamisesta ja käytöstä aiheutuvat eettiset riskit. Nämä riskit voidaan jakaa välittömiin, keskipitkän ja pitkän aikavälin huoliin. Koska käsittelemme nopeasti kehittyvää alaa, ei näihin ole valmista vastausta, vaan valtaosa tekemistämme päätöksistä tulee tapahtumaan kokeilun ja erehdyksen kautta. Tästä syystä on valitettavasti luultavaa, että osa, ellei jopa valtaosa meistä tulee käyttämään tekoälyä sellaisiin tarkoituksiin, joilla on negatiiviset seuraukset joko lyhyellä tai pitkällä aikavälillä. (Silberg & Manyika, 2019.)



Kun tiedetään miten tekoälytyökalut toimivat, että ne linjaavat vahvoja eettisiä ja akateemisia standardeja ja että niillä tuotetaan aidosti laadukasta jälkeä, niistä voi olla meille merkittävä hyöty. Tätä varten tulisi rakentaa eettinen viitekehys, jonka mukaan toimia, ja jonka perusteella valita käyttöön sopivimmat työkalut tietäen, että niiden data on vinoutumatonta. Tämä ei kuitenkaan ole kovin yksinkertaista tai edes mahdollista. Jo pelkästään uusia sovelluksia ja tekoälytyökaluja julkaistaan päivittäin sadoittain. Kuinka pystytään varmistamaan niiden toimivuus, eettisyys ja laatu?

Open AI:n marraskuussa 2022 julkaisema ChatGPT sai nopeasti huomiota ja saavutti 100 miljoonan aktiivisen käyttäjäkunnan jo tammikuussa 2023. Vertailun vuoksi TikTok saavutti samaisen rajapyykin 9 kuukaudessa ja Instagram 2,5 vuodessa. (Patel & Ahmad, 2023.) Instagramin julkaistua Threads-sovelluksensa vastineeksi Twitterille heinäkuussa 2023, saavutti se ensimmäiset kaksi miljoonaa käyttäjää muutamassa tunnissa ja 100 miljoonan käyttäjän rajan vain muutamassa päivässä (Peters & Porter, 2023.).

Sam Altmanin mukaan ChatGPT:lla voi olla monia negatiivisia seurauksia: se voi lisätä ennakkoluuloja ja disinformaation leviämistä tai kehittää koneita, jotka ovat yhteiskuntaa älykkäämpiä. Näiden lisäksi sillä saattaa

Altmanin mukaan olla muita sellaisia seurauksia, jotka ovat niin kamalia, ettemme voi edes kuvitella millaisia ne voisivat olla. Se mikä Altmania kuitenkin pelottaa kaikista eniten on muutoksen nopeus. Tämän tiedon valossa, tuntuu uskomattomalta, että tämä teknologia ylipäättään kehitettiin. Miksi OpenAI kehitti ChatGPT:n ja toi sen kaikkien saataville ennen kuin varmuutta sen käytön turvallisuudesta oli? Sam Altmanin mukaan, ChatGPT on suurin ihmisen tähän mennessä kehittämä teknologia, jolla luonnollisesti on potentiaalinen kääntöpuoli. Altmanin mukaan, olisi luultavasti ollut suurempi riski kehittää ChatGPT kokonaan laboratorio-olosuhteissa ja julkaista se "valmiina" kaikkien käyttöön. Altman uskoo, että ChatGPT pystyy myös kaikkien mahdollisten kauhuskenaarioiden vastaakohtaan. (Jarvis 2023.)

Mira Murati, on yksi ChatGPT:n kehittäjistä OpenAI:lla. Uusia teknologioita kehittäessä, ei voi Muratin mukaan etukäteen tietää kaikkia niiden kyvykkyyksiä, potentiaalisia vaikutuksia tai rajoitteita. Hänen mukaansa OpenAI:n tavoitteena on taata, että he pystyvät ennustamaan ChatGPT:n kykyjä ja rajoitteita. Murati sanoo, että näin tehokkaan teknologian rakentamiseen sisältyy aina myös suuri riski. (Jarvis 2023.). Osaamme siis rakentaa asioita, mutta emme tarkalleen tiedä, miten ne toimivat. Ovatko insinööritaidot menneet tieteen edelle?

Tekoäly on uskomattoman älykäs, mutta samalla järkyttävän typerä. Kysyttäessä ChatGPT 3.5:ltä (7/2023) kauanko kestää 30 paidan ulkona kuivattamiseen, kun viiden paidan kuivumiseen menee viisi tuntia, tekoäly vastaa $30 \times 1 \text{ h} = 30 \text{ h}$. Tekoäly toki oppisi pikkuhiljaa laskemaan tehtävän, jos syöttäisimme sille opetusdatana monia edeltävän kaltaisia laskutehtäviä. Nick Bostrom on ruotsalainen filosofi, joka työskentelee Oxfordin yliopistossa. Hän on esitellyt kuuluisan ajatusleikin "The paperclip maximizer problem", jossa tekoälyä pyydetään valmistamaan mahdollisimman paljon paperiliittimiä. Loppujen lopuksi tekoäly tappaa ihmisetkin liittimien raaka-aineeksi. Voisimme parantaa tekoälylle anta-
maamme kehoitetta alleviivaten, ettei se saisi tappaa ihmisiä toteuttaessaan tehtävänsä, mutta se voisi sen sijaan tappaa kaiken kasvillisuuden, jolloin lopputulema olisi sama. On lukuisia muitakin seikkoja, jotka tekoälyn tulisi ottaa huomioon maksimoidessaan paperiliittimien määrää – muistaisimmeko mainita nämä kaikki kehoitteessa? Tekoälyltä puuttuu meille ihmisille ominainen terve järki ja ymmärrys siitä, miten maailma toimii. Terve järki onkin ollut jo pitkään tekoälykentällä suuri haaste, koska sen koneellinen saavuttaminen on osoittautunut olevan varsin hankalaa. Vaikka tekoälyn saralla onkin saavutettu suuria harppauksia viime aikoina, se on edelleen melko heikko kompuroidessaan sellaisissa tehtävissä, joista pienet lapsetkin selviytyisivät. (Choi, 2023.)

älykäs mutta typerä

“Whether a given sentence sounds plausible does not guarantee its factuality.”

James Vincent, The Verge

Koneoppimismalleja kritisoidaan joskus siitä, että niiden toiminta vaikuttaa mustalta laatikolta, joka ei kerro, mihin sen päätös perustuu. Todellisuudessa asia ei ole näin mustavalkoinen. Hyvin yksinkertaisten mallien toiminta on suoraviivaista. Joissakin malleissa on myös sisäänrakennettuja mekanismeja, jotka kertovat, mihin niiden toiminta perustuu. Mustanlaatikon ongelma tulee esille usein syviä neuroverkkoja tarkastellessa, koska niiden kouluttamiseen käytetty datan määrä on valtava ja säädettävien parametrien määrä voi kohota satoihin miljooniin. Näidenkin mallien toimintaa voidaan kuitenkin tutkia erilaisilla menetelmillä. (Hursti ym. 2023a.)

Data

Karkeasti katsottuna tekoälyä voidaan harjoittaa kolmen tyyppisellä datalla. Vapaasti saatavissa oleva (internet)data sisältää yleensä paljon rasismia, seksismiä ja disinformaatiota. Vapaasti saatavissa olevan datan lisäksi tekoälyä voidaan kouluttaa myös tekoälyn kouluttamiseen erikseen tehdyllä datalla sekä ihmisten palautteella. Kun tieto koulutusdatasta ei ole avointa, meidän on mahdoton tietää mitä tekoälylle on opetettu. Kahden viimeisen opetusmenetelmän kustannukset ovat suuret ja lisäksi on uutisoitu monista eettisistä ongelmista esimerkiksi juuri OpenAI:n ihmispalautetyöntekijöiden työoloista. (Choi, 2023.) Täydellisessä maailmassa data olisi vinoutumatonta, siihen olisi sisällytetty erilaista kulttuurista kontekstia ja se huomioisi yhdenvertaisuuden, diversiteetin ja inklusion. (World Economic Forum 2023c.)

Keskiössä tällä hetkellä ovat sovellukset, joista monet pohjautuvat suuriin datamääriin. Onkin syytä tarkastella kriittisesti datan oikeellisuutta ja kattavuutta, sekä datan käyttöön liittyviä eettisiä ja yksityisyydensuojaa koskevia seikkoja. On tärkeää muistaa, että tekoäly on huono tekemään sellaisia päätelmiä, joihin sitä ei ole koulutettu. Tällaisissa tapauksissa sen päätökset voivat olla huonoja tai jopa vääriä. (Hursti ym. 2023a.)

Oikeudenmukaisuus ja vinoumat

Miten oikeudenmukaisuus määritellään? Onko se sitä, että samanlaisia yksilöitä kohdellaan samalla tavalla? Vai että tulokset ovat oikeudenmukaiset ryhmien kesken, erityisesti potentiaalisesti haavoittuvien ryhmien välillä. Yleisen määritelmän tai mittarin luominen oikeudenmukaisuudelle ei todennäköisesti koskaan tule olemaan mahdollista. Sen sijaan erilaisia mittareita ja standardeja, sekä näiden yhdistelmiä tarvitaan mahdollisimman oikeudenmukaiseen tuloksen saavuttamiseksi. (Silberg & Manyika, 2019.)

Mitä oikeudenmukaisuus on?

On selvää, että mitä laajemmin tekoälyä otetaan käyttöön herättää se keskustelua tekoällyn vinoumista ja käytön oikeudenmukaisuudesta. Tiedetään, että tekoällyn antamat vastaukset voivat olla hyvinkin vinoutuneita ja harhaisia. Toisaalta ihmisten tekemät päätökset voivat olla yhtä lailla virheellisiä ja ne ovat lisäksi usein vielä vaikeasti tutkittavissa tai tarkasteltavissa, koska ihmiset eivät ehkä aina ymmärrä, mitkä tekijät vaikuttavat heidän päätöksentekoonsa. (Silberg & Manyika, 2019.)

Tekoällyn myötä kuvien, äänien ja videoiden manipuloiminen on viety aivan uudelle tasolle. Meidän on jatkuvasti vaikeampi erottaa totuutta kuvitelmasta tai meille esitetystä harhasta. Sama pätee myös tekstiin. Medialukutaito nousee äärimmäisen tärkeään asemaan ja sen tärkeyttä tulisi korostaa sekä kouluissa että työpaikoilla. Tulisi muistaa, että mitä enemmän kielimallia käyttää, sen hallusinoinnille tulee ikään kuin sokeaksi. ChatGPT on kehitetty perustelukoneeksi, ei faktatietokannaksi. Näitä järjestelmiä koulutetaan perustelemaan ja laskemaan todennäköisyyksiä, ei muistamaan tai kertomaan totuuksia. (Jarvis, 2023.)

median- ja tekoällynlukutaito

Mikäli koulutusdata tai mallien kehittäjien tekemät päätökset ovat epätäydellisiä, voivat mallit tuottaa algoritmista vinoumaa. (Chui, Roberts ym, 2023.). Tekoälymallien sisältämät ihmisten ja yhteiskunnan ennakkoluulot ja vinoumat ovat usein algoritmin sijaan opetusdatassa. Algoritmit voivat kuitenkin oppia epähyväksyttäviä tai jopa laittomia tilastollisia korrelaatioita. Esimerkiksi vakuutusmallin toimintaa voidaan pitää laittomana ikään perustuvana syrjintänä, mikäli se ei myönnä vakuutusta vakuutusnottajan korkean iän vuoksi. Tätä kutsutaan algoritmiseksi syrjinnäksi. Algoritmisen syrjintä juontaa juurensa ihmisen tekemään syrjintään ja tästä syystä on erityisen tärkeää tietää, minkälaista opetusdataa opettamiseen on käytetty tai jos sitä on mahdotonta selvittää, tulisi ainakin tekoällyn antamaan tulokseen suhtautua varauksella ja olla tietoinen algoritmisen syrjinnän mahdollisuudesta. Toisin kuin ihmisten tekemät päätökset, tekoälyllä tehdyt päätökset voidaan kuitenkin periaatteessa tutkia ja kyseenalaistaa. (Silberg & Manyika, 2019; Elements of AI, 2023f.) Pahentaako tekoäly jo olemassa olevia päätöksenteon vinoumia vai tekeekö se päätöksenteon prosessista luotettavamman?

Pahentaako vai parantaako?

Kaikkien yritysten, joilla on toimintaa tai asiakkaita EU:n alueella ovat velvoitettuja toimimaan EU:n yleisen tietosuoja-asetuksen, GDPR:n mukaisesti. Tämä lisää merkittävästi avoimuutta ja läpinäkyvyyttä. Algoritmit ovat kuitenkin monesti yritysten liikesalaisuuksia ja vaikka algoritminen toiminta olisi täysin julkista tietoa, olisi silti hyvin vaikeaa tunnistaa mikä kohta algoritmia tai sen datasta johtaa syrjintään. Lisäksi asetusten tulkinanvaraisuus asettaa omat haasteensa kattavien selvitysten saamiseen. Tekoälyn yhdistäminen datankeruuseen muodostaa uusia yksityisydensuojaan liittyviä haasteita. Moniin yksityisydensuojan loukkauksiin voitaisiin ehkä puuttua ajoissa, mutta niiden ennustaminen on vaikeaa. (Elements of AI, 2023f.)

Ovatko ihmisten tekemät päätökset parempia?

On tärkeää miettiä, missä tilanteissa tarvitaan ihmisen harkintaa ja missä tilanteissa täysin automatisoitu päätöksenteko on sallittua. Kuka päättää, milloin tekoälyjärjestelmä on minimoitu vinoumilta, ja että se voidaan turvallisesti ottaa käyttöön? Miten määritellään milloin ihmisten tekemät päätökset ovat oikeudenmukaisia tai vääristymättömiä? Jos algoritmisen päätöksenteon arvioiminen on hankalaa, niin ihmisten tekemien päätösten arvioiminen on lähes mahdotonta. Yleisesti hyväksymme päätökset, joiden katsomme perustuvan oikeudenmukaiseen prosessiin tai jos ajattelemme päätöksiä tekevällä ryhmällä olevan tarpeeksi monimuotoisia näkökulmia. Onko menettelyllinen oikeudenmukaisuus sama kuin lopputuloksellinen oikeudenmukaisuus? Samalla kun keskustelemme automaattisen päätöksenteon oikeudenmukaisuudesta, pitäisikö meidän kohottaa myös ihmisten tekemän päätöksenteon oikeudenmukisuuden rimaa? (Silberg & Manyika, 2019.)

tietoisuus!

Vääristymien minimoiminen on tärkeä edellytys sille, että ihmiset voivat luottaa tekoälyjärjestelmiin. Tätä varten tulisi olla keskusteleva, avoin ja läpinäkyvä, mutta ennen kaikkea tietoinen niistä vinoumista, joita tekoälyn käyttöön liittyy. Alan nopeasti kehittyessä ja kokemuksen kasvaessa on olennaista jatkuvasti, kriittisesti ja monitieteellisesti tarkastella ja arvioida tekoälyprosessin päätöksentekoa. On myös olennaista, että tekoälyala itsessään kuvastaisi yhteiskunnan monimuotoisuutta. (Silberg & Manyika, 2019.)

Hinta ja vallan keskittyminen

Viime vuosikymmenen aikana tekoälyn kehityksen selkeä trendi on suuntautunut yhä suurempiin mittakaavoihin. Tämä on luonut riippuvuutta massiivisiin datamääriin ja suuren datan määrän käsittelyyn tarvittavan laskentatehoon. Nämä riippuvuudet antavat merkittävän etulyöntiaseman pienelle joukolle teknologiayrityksiä, jotka hallitsevat näitä resursseja ja jotka operoivat myös laajalle levinneitä alustaympäristöjä. Näillä yrityksillä on täten myös etulyöntiasema suuren mittakaavan tekoälytoteutuksissa. Tämä infrastruktuurinen ja taloudellinen valta haittaa kilpailua tekoälyalalla, koska pienet yritykset kohtaavat merkittäviä markkinoille pääsyn esteitä suurten yritysten dominoidessa alaa. Valta-aseman keskittyminen tuottaa myös kuluttajille vahinkoa. Tämä saattaa esimerkiksi vaarantaa kuluttajan yksityisyyttä ja tietoturvaa, edistää väärin ja harhaanjohtavien tietojen leviämistä sekä eriarvostaa ja syrjiä. (Myers West & Vipra, 2023.)

Ensimmäisten perusmallien kehittäminen vaati merkittäviä sijoituksia, koska niiden kouluttamiseen tarvittiin huomattavaa laskentatehoa ja ihmistyötä mallien hienosäätämiseksi. Tämän seurauksena ne kehittyivät pääasiassa muutamien suurten teknologiayritysten, merkittävän sijoituksen saaneiden start-up-yritysten ja joidenkin avoimen lähdekoodin tutkimuskollektiivien (kuten BigScience) toimesta. Pienempien mallien tehokkaamman koulutuksen parissa on tehty viime aikoina paljon työtä, ja tämä onkin avannut markkinat myös uusille tulokkaille. Monet start-up-yritykset ovat jo onnistuneet kehittämään omia mallejaan – esimerkiksi Cohere ja Anthropic rakentavat ja kouluttavat omia suuria kielimallejaan. (Chui ym. 2023.)

SemiAnalysis:n pääanalytiikko Dylan Patelin helmikuussa 2023 julkaistun raportin mukaan ChatGPT:n ylläpito ja käyttö GPT-3-mallilla maksaa noin 700 000 USD päivässä. Tämän perusteella jokaiselle sille asetetulle kehotteelle tulisi hintaa 0,36 \$. Korkeat kustannukset johtuvat pääasiassa kalliista palvelimista ja niiden ylläpitokustannuksista. Vertailun vuoksi yhden Google-haun hinnaksi on arvioitu 0,0106 \$. Haastattelussa Alphabetin (Google) John Hennessy kertoi Reutersille, että tavalliseen Google-hakuun verrattuna Bard-haku maksaa 10 kertaa enemmän, vaikkakin mallien hienosäätö tulee vähentämään kustannuksia nopeasti. (Dastin & Nellis, 2023.). (Patel & Ahmad, 2023.)

chatgpt kehote: 0,36 €/kpl
google-haku: 0,0106 €/kpl

Alan ympäristövaikutukset ovat mittavat: laskennallisesti intensiivinen tekoäly tarvitsee toimiakseen datakeskuksen. Datakeskuksilla on korkeat energiakustannukset ja valtava hiilijalanjälki. (Myers West & Vipra, 2023.) Sijainnista ja energianlähteestä riippuen datakeskukset tarvitsevat valtavia määriä vettä esimerkiksi sähköntuotantoon sekä palvelimien jäähdyttämiseksi ja ylikuumenemisen estämiseksi. Vesimäärän ja ylipäätään energiankulutuksen todentaminen on vaikeaa, sillä monet yritykset eivät raportoi käytöstään läpinäkyvästi. Suuret teknologiayritykset ovat paineen alla asettaneet julkisia, kunnianhimoisia kestävyystavoitteita, mutta näiden on arvosteltu olevan enemmän markkinointia kuin todellista toimintaa. (Myers West & Kak, 2023, 100–101.) Tekoälysovelluksien energiankulutuksen arvioidaan tuplautuvan 3–4 kuukauden välein. (Hursti, Puuppo, Raita-Hakola, Riihioho, Lind, Hämäläinen, Kettunen, Uotinen, Waselius, & Pölönen, 2023a.)

Ympäristövaikutukset

Tekoälymallien kouluttaminen on generatiivisen tekoälyn energiantensiivisin vaihe. Eri laskelmien mukaan yhden suuren kielimallin kouluttamiseen vaaditaan satojen tuntien edestä laskentatehoa. Datakeskukset sijaitsevat ympäri maailmaa ja riippuen niiden käyttämisestä energianlähteistä saattaa datakeskusten hiilijalanjäljet vaihdella huomattavasti. Eräissä tutkimuksissa vertailtiin GPT-3-mallin koulutuksen hiilijalanjälkeä kolmen eri palveluntarjoajan datakeskuksessa ympäri maailman (Google GCP, Amazon AWS ja Microsoft Azure). Alhaisimmillaan mallin kouluttaminen tuotti noin 4 000 kg hiilidioksidia. Näin pienet päästöt mahdollistui Kanadassa alueilla, joilla vesivoima on pääenergiälähde. Kolme datakeskusta, joissa kouluttamisesta aiheutui suurimmat päästöt sijaitsivat Etelä-Afrikassa ja Intiassa. Molemmissa maissa sähköverkot perustuvat pitkälti kivihiileen ja öljyyn. Kaikista suurimmat päästöt oli Etelä-Afrikkalaisessa datakeskuksessa, jossa mallin koulutus tuotti lähes 211 000 kg hiilidioksidia. Tietolähteestä riippuen, tekoälyn ympäristövaikutukset vaihtelevat merkittävästi. (Shrink That Footprint, 2022.)



Tutkijoiden eri arvioiden mukaan GPT-3:n kouluttaminen olisi siis suurimillaan tuottanut noin 200 tonnia hiilidioksidia. Keskimääräinen amerikkalainen tuottaa noin 15 tonnia hiilidioksidia vuodessa, keski-vertohenkilö noin 4–5 tonnia vuodessa. Pienimmillään kouluttaminen tuotti noin 4 tonnia hiilidioksidipäästöjä, joka vastaa noin neljäsosaa keski-vertamerikkalaisen vuosittaisesta hiilidioksidituotannosta tai yhden keski-vertohenkilönvuosittaista hiilidioksidituotantoa. Vastaavasti, tällä menetelmällä kouluttaen GPT-3 päästöt vastaisivat karkeasti yhden auton vuosittaisia hiilidioksidipäästöjä tai mikäli koulutus tapahtuisi suurimpien päästöjen datakeskuksissa, 40 auton vuosittaisia hiilidioksidipäästöjä. (Shrink That Footprint, 2022.)

kaikki on suhteellista



Kielimallien koulutuksella on vaikutus maailman hiilidioksidipäästöihin, mutta vielä GPT-3:n aikaan ne olivat suhteellisen pieniä muihin saastuttaviin teollisuudenaloihin verrattuna. Suurten kielimallien aiheuttamat hiilidioksidipäästöt olivat tuhat kertaa pienemmät kuin kolmen saastuttavimman eli ruokateollisuuden, liikenteen, sekä lämmityksen ja jäähdytyksen. Tilanne voi kuitenkin muuttua kielimallien kasvaessa ja mikäli koulutuksen tehokkuus ei muutu. (Shrink That Footprint, 2022.)

Viimeisimmässä ympäristöraportissaan Microsoft ilmoitti, että sen maailmanlaajuinen vedenkulutus kasvoi 34 % vuodesta 2021 vuoteen 2022. Samalta ajanjaksolta Google ilmoitti 20 %:n kasvusta vedenkulutuksessa. Tämän huomattavan nousun ulkopuoliset tutkijat yhdistivät liittyvän tekoälytutkimukseen. Erään arvion mukaan ChatGPT kuluttaa 500 ml vettä jokaista 5–50 sarjan kyselyä kohden. Vedenkulutuksen suuruus vaihtelee palvelimen sijainnista ja vuodenajasta riippuen. Arvio sisältää suoran vedenkulutuksen lisäksi myös epäsuoran vedenkulutuksen, kuten datakeskuksille sähköä tuottavien voimalaitosten jäähdyttämiseen käytetyn veden. Epäsuoraa vedenkulutusta yritykset eivät yleensä mittaa tai huomioi raporteissaan. (Microsoft 2023, 26; O'Brien & Fingerhut, 2023)

Tekijänoikeus, yksityisyys ja turvallisuus

Koulutusdata ja mallien tuotokset voivat aiheuttaa merkittäviä immateriaalioikeusriskejä, esimerkiksi tekijänoikeus-, tavaramerkki-, patentti- tai muita laillisesti suojattujen materiaalien loukkauksia. Tekoälyä hyödyntäessään organisaatioiden on ymmärrettävä, mitä dataa mallin koulutukseen on käytetty ja miten se voi ilmetä työkalun tuotoksissa. (Chui, Roberts ym, 2023.) Yhdysvalloissa käynnissä olevilla tekijänoikeusoikeudenkäynneillä on merkittäviä vaikutuksia generatiivisten tekoälyjärjestelmien tulevaisuuteen. Jos kantajat voittavat, ainoastaan sellaiset generatiiviset tekoälyjärjestelmät voivat olla Yhdysvalloissa laillisia, jotka on koulutettu julkisesti saatavilla olevilla teoksilla tai sellaisella aineistolla, johon on asianmukaiset lisenssit. Tämä vaikuttaisi kaikkiin, jotka hyödyntävät generatiivista tekoälyä, integroivat sen tuotteisiinsa ja käyttävät sitä tieteelliseen tutkimukseen.

Yksityisyysongelmia voi ilmetä, jos käyttäjät antavat tekoälymalleille tietoja, jotka myöhemmin päätyvät mallien tuotoksiin tai koulutusdataan sellaisella tavalla, josta käyttäjät voidaan tunnistaa. Generatiivista tekoälyä voidaan myös käyttää haitallisen sisällön, kuten disinformaation, syväväärennösten ja vihapuheen luomiseen ja levittämiseen. (Chui, Roberts

ym, 2023.) Monet tekoälyn mukana tuomista haasteista eivät sinänsä ole uusia. Misinformaatiota ja väärennöksiä on ollut olemassa jo ennen painettua tekstiäkin, mutta tekeminen ja levittäminen on teknologian myötä tullut helpommaksi. Tekoäly helpottaa huijauksien läpiviemistä entisestään. Ihmiset ovat kuitenkin oppineet vuosien mittaan, ettei sähköpostilla lähestyvälle ghanalaisprinssille kannata lähettää pankkikorttitietojaan. Pahantahtoiset toimijat voivat hyödyntää generatiivista tekoälyä nopeuttaakseen ja kehittääkseen kyberhyökkäyksiä. Sitä voidaan myös manipuloida tuottamaan haitallisia tuloksia. Esimerkiksi tekniikalla nimeltä prompt injection, kolmas osapuoli antaa mallille uusia ohjeita, jotka huijaukseen mallia tuottamaan sellaisia tuotoksia, jotka eroavat mallin tuottajien ohjeista eikä ole tarkoitettu loppukäyttäjälle. (Chui, Roberts ym, 2023.)

Teknologia kehittyä, huijaukset paranevat, mutta samalla huijauksien kohteet sivistyvät, nahka paksuuntuu eikä huijaaminen ole enää niin helppoa. Näin voisi olettaa käyvän myös tekoälyn kanssa. Sitä mukaan, kun tekoälyllä kehitellään uusia huijauksia ja syvävärennöksiä, syvävärennösten tunnistaminen helpottuu tekoälyn avulla. Tekoälyä voidaan käyttää sekä hyviin että huonoihin tarkoituksiin sen kaikilla saroilla syvävärennöksistä kyberturvallisuuteen.

Työ on ollut jatkuvan murroksen alla koko ihmiskunnan historian ajan. Uudet keksinnöt ja teolliset uudistukset sekä läpimurrot ovat kautta aikain muuttaneet työnteon tarve ei kuitenkaan ole missään vaiheessa kadonnut kokonaan vaan ainoastaan muuttanut muotoaan. Uudet työt ovat kautta historian olleet turvallisempia, vaihtelevampia ja luovempia entisiin verrattuna ja ihmisten hyvinvointi on pääsääntöisesti kasvanut. Ongelmaksi ei tälläkään kertaa, tekoälyn murroksessa tule muodostumaan työn loppuminen, vaan muutoksen nopeus. Emme ehdi kouluttaa työvoimaa uusiin työpaikkoihin tarpeen nopeasti kasvaessa. Koulutus ja elinikäinen oppiminen tulevat olemaan tulevaisuudessa jokaisen työntekijän välttämättömyyksiä. (Elements of AI 2023f)

Generatiivista tekoälyä voidaan käyttää automatisoimaan, täydentämään ja nopeuttamaan työtä. Sen sijaan, että kauhistelisimme sen mahdollisuutta korvata ihminen eri työtehtävissä, meidän kannattaisi keskittyä tapoihin, joilla generatiivinen tekoäly voi tehostaa työtä. Ongelmaksi tulee muodostumaan se, että sen vaikutukset voivat olla epätasaisesti negatiivisia tietyille ryhmille ja yhteisöille (Chui, Roberts ym, 2023.).

”Kun tekoäly omaksuu ihmiskielen ja hallitsee sen, sillä on kaikki tarvittava koteloidakseen meidät illuusioiden maailmaan. Tuhansien vuosien ajan profeetat, runoilijat ja poliitikot ovat käyttäneet kieltä ja tarinankerrontaa manipuloidakseen, hallitakseen ihmisiä ja muokatakseen yhteiskuntaa. Tekoäly todennäköisesti pystyy siihen tulevaisuudessa ja kun se pystyy, sen ei tarvitse lähettää tappavia robotteja ampumaan meitä. Se voi saada ihmiset vetämään liipaisimesta, jos se todella niin haluaa.” Näin sanoo Yuval Noah Harari, historioitsija, filosofi ja kirjailija, joka on varoittanut ”ihmishakeroinnin vaaroista” ja ”hyödyttömän luokan noususta” jo vuodesta 2014. Hän uskoo, että tekoälyllä on paljon positiivista potentiaalia, mutta loi nämä provosoivat ilmaisut nimenomaan kiinnittääkseen ihmisten huomion tekoälyn vaaroihin. (Humanity is not that simple, 2023; Yuval Noah Harari, 2022.)

Vaikutus työvoimaan

koulutus ja elinikäinen oppiminen

Polarisoituminen ja vaikutukset demokratiaan

“The same capabilities that make AI dangerous in the hands of bad guys with bad goals make it powerful in the hands of good guys with good goals – specifically the good guys whose job it is to prevent bad things from happening.”

Marc Andreessen,
software engineer and entrepreneur

Hararin sanoo, että sosiaalisen median algoritmien myötä elämme maailmassa, jossa voimme harhautua luulemaan, että se mitä näemme, on totuus. Sosiaalinen media on lisännyt sosiaalista polarisoitumista, vaikuttanut mielenterveyteemme ja jopa epävakauttanut demokraattisia yhteiskuntia. Miljoonien ihmisten on vaikea erottaa todellisuutta illuusiosta. Sosiaalisen median melko alkeelliset algoritmit ovat onnistuneet luomaan sosiaalista ja poliittista kaaosta. Miten käy, kun uudet, tehokkaammat algoritmit korvaavat edeltäjänsä? Hararin mukaan meillä on käsissämme uusi massatuhoase, joka voi tuhota sosiaalisen ja mentaalisen maailmamme. Harari muistuttaa, että vuodesta 1945 asti olemme olleet tietoisia siitä, että ydinvoima voi tuhota ihmissivilisaation, mutta se voi myös auttaa meitä tuottamaan halpaa energiaa. Siksi varmistammekin, että ydinteknologiaa käytetään ensisijaisesti hyvään tarkoitukseen. Yksi ero ydinaseiden ja tekoälyn välillä on Hararin mukaan se, etteivät ydinaseet voi tuottaa tehokkaampia ydinaseita. Tästä syystä tekoälyä pitäisi Hararin mukaan säännellä ja valvoa, ennen kuin se sääntelee meitä. Onko tekoäly enää hallinnassamme, kun ihmiset eivät varmuudella osaa sanoa, miten tekoäly tulee toimimaan? (Humanity is not that simple, 2023.)

Tekoäly on ensimmäinen teknologia, joka voi tehdä itsenäisiä päätöksiä jopa meistä ihmisistä. Se on enenevässä määrin mukana päätöksenteossa. Toki olemme opettaneet tekoälyjärjestelmälle, miten toimia, mutta on selvää, että osa vallastamme esimerkiksi juuri päätöksenteossa on siirtynyt tekoälylle. Tekoäly osaa myös tuottaa ideoita. Harari peräänkuuluttaa tekoälyn kehityksen mahdollista rajoittamista, mutta tärkeämpänä hän pitää tekoälyn käyttöönoton rajoittamista yhteiskunnassa. (Humanity is not that simple, 2023.)

kehityksen ja
käyttöönoton rajoittaminen

Harari sanoo, että helpoin tapa vaikuttaa ihmisen tekemiin päätöksiin, poliittisiin näkökulmiin tai esimerkiksi ostamaan tuote on luoda läheinen yhteys ihmiseen. Jos sallimme algoritmien oppia luomaan meihin tällaisia yhteyksiä, tulee se tuhoamaan demokratian, koska demokratia perustuu ihmisten väliselle keskustelulle ja kielelle. Jos keskustelu siirtyy ihmisen ja tekoälyn väliseksi, tietää se Hararin ennusteiden mukaan demokratian loppua. Ongelmallista on se, että on tilanteita, joissa käyttäjä ei välttämättä tiedä olevansa tekemisissä koneen kanssa. Salaliittoteorioiden ja väärin uutisten levittäminen on helpottunut algoritmien opittua, että juuri vihapuhe ja skandaalit ovat paras tapa kiinnittää ihmisten huomio. Meillä ei ole tietoa siitä, millaiset sosiaaliset ja psykologiset seuraukset tällä teknologialla tulee olemaan. Mitä tapahtuu, kun algoritmit oppivat, että läheisen yhteyden luominen ihmiseen on paras tapa vaikuttaa tämän päätöksentekoon? (Humanity is not that simple, 2023.)

demokratia perustuu ihmisten
väliselle keskustelulle ja kielelle

Hararin mielestä koneen, mallin, alustan tai muun teknologian jäljittelessä tai "väärentäessä" ihmistä, siitä pitäisi rangaista samalla tavalla kuin rahan väärentämisestä rangaistaan. Rahan väärentäminen on helppoa, kuten tekoälyn avulla ihmistenkin, mutta mikäli se olisi sallittua koko systeemimme kaatuisi. Samaa menettelytapaa tulisi Hararin mielestä soveltaa tekoälyyn. (Humanity is not that simple, 2023.)

**“It’s the biggest
experiment in
history conducted
on billions
of people.”**

Yuval Noah Harari, historian,
philosopher and author

Onko meillä tulevaisuudessa kykyä sellaiseen kriittiseen ajatteluun, että pystymme epäilemään tekoälyltä saamaamme vastausta ja etsimään vaihtoehtoisia, omia vastauksia? Hakukoneiden, kuten Googlen kanssa on voinut törmätä vastaavanlaiseen ongelmaan, mikäli on hakutuloksista valinnut ensimmäisen tuloksen, joka on kärkisijoille noussut juuri algoritmien saattamana. Toisaalta hakukoneen antaman vastauksen on tähän asti joutunut lukemaan itse ja vasta sitten tekemään siitä omat johtopäätökset. Tekoälyn myötä riittää, että kysyy kysymyksen. (Humanity is not that simple, 2023.)

suurin uhka
ihmiskunnalle

Tekoälyn suurin uhka ihmiskunnalle on se, että huomiomme keskittyisi liiaksi siihen ja jättäisimme puuttumatta todellisempiin uhkiin kuten demokratian vastavoimiin, ydinaseisiin, ilmastonmuutokseen ja luonnonkatastrofeihin. (Elements of AI 2023f).

Seuraavat tutkinnon saavat opiskelijat ovat ensimmäinen sukupolvi, jonka koulutuksessa on käytetty generatiivista tekoälyä. Sen käytöstä uskotaan olevan paljon hyötyä, mutta myös potentiaalia haitallisiin, jopa vaarallisiin seurauksiin. Generatiivisella tekoälyllä alkaa kasvavassa määrin olemaan vaikutusta myös opetuksen suurien kielimallien, kuten ChatGPT muuttamalla totuttua oppimisympäristöämme. Jotkut näkevät sen ihannetyökaluna huijaamiseen ja uhkana koko opetukselle, kun toiset taas hyödyntävät sitä arvokkaana henkilökohtaisena avustajana, joka tulevaisuudessa voisi tunnistaa opiskelijan kiinnostuksen kohteet ja oppimistyylin, räätälöidä motivoivaa sisältöä, mitata opiskelijan oppimista, huomata milloin opiskelijan motivaatio katoaa, ja antaa välitöntä palautetta. Miten tekoäly omaksutaan oppilaitoksissa? Instituutiot, kuten yliopistot tulevat luultavasti yrityksiä hitaammin vapauttamaan käyttöönsä tekoälyn koko potentiaalilin, koska niitä ei ole rakennettu nopeaan muutosten implementointiin ja päätöksentekoon. Pandemia kuitenkin osoitti, että muutokseen pystytään nopeallakin aikataululla, mikäli siihen on halu ja tarve. (World Economic Forum 2023 c&d.)

muutokseen pystytään,
mikäli siihen on halu ja tarve

Tekoäly ei vain muuta tapaa, jolla elämme, työskentelemme ja opimme, vaan myös tapaa jolla opetamme ja opiskelemme. Sillä on potentiaalia parantaa korkeakoulutuksen laatua, saavutettavuutta ja tehokkuutta sekä luoda uusia mahdollisuuksia innovaatioille ja yhteistyölle. Katoaako kriittinen ajattelu, luovuus, tuottavuus ja yrittäjähenki tekoälyn myötä? Miten tekoälyn eettinen käyttö varmistetaan? Kysymyksiä tuntuu olevan enemmän kuin vastauksia. Jakamalla ymmärrystä tekoälystä luomme samalla pohjan hyväksynnälle. On tärkeää, että kaikilla on perusymmärrys siitä, kuinka tekoäly toimii, jotta voimme ymmärtää sen rajoja ja mahdollisuuksia. Koska emme voi kieltää tekoälytyökalujen olemassaoloa tai sulkea niitä opetuksen ulkopuolelle, meidän olisi parasta ottaa ne käyttöön tiedostaen niiden rajoitukset ja laatia tapoja niiden hyödyntämiseen opetuksen edistämiseksi. Tekoälyn kehittyessä ja tehtävien automatisoituessa on hyvä miettiä, mitkä tehtävät jäävät automaation ulkopuolelle. Tulevaisuuden opetuksessa empatia, luovuus, uteliaisuus ja kriittinen ajattelu tulee korostumaan.

Opiskelun tarkoituksena ei suinkaan ole tehtävien tekeminen, vaan oppiminen, ajatteleminen ja opitun tiedon soveltaminen käytäntöön. Luultavasti tekoäly tulee vaikuttamaan kykyymme päätellä, koska meille tarjotaan vastauksia kuin buffetpöydästä. Navigaattoreiden yleistyminen on esimerkiksi vaikuttanut kartanluku- ja suunnistustaitoihimme. Ei ole poikkeuksellista, ettei löydä kotikaupungissaankaan määränpäähän ilman koko ajan päivittyvää karttanäkymää ja ääniohjausta. Voidaan toki miettiä, onko meidän enää tarpeen osata suunnistaa. Teknologian kehittyessä ja sen käytön yleistyessä arkielämän taidot ovat merkittävästi muuttuneet. Tekoälyn suurin uhka opiskelijalle on kuitenkin se, että opiskelija lakkaa ajattelemasta itse.

suurin uhka opiskelijalle

“We need to go faster because we will continue to be behind.”

Kim Majerus, Vice-President,
Global Education, Amazon

Suuret kielimallit eivät ole faktapankkeja, mutta ne tekevät esimerkiksi kirjoittamisesta nopeampaa ja helpompaa monelle, jolle se muuten on haasteellista. Näiden työkalujen käyttö nimenomaan kirjoituksen tukena ennemmin kuin tiedonlähteenä olisi niin opettajien kuin oppilaidenkin näkökulmasta tehokasta. Erilaisten oppimisvaikeuksien kanssa kamppailevien oppilaiden oppimista nämä työkalut voivat helpottaa suuresti. Erityisen paljon hyötyä suuriin kielimalleihin pohjautuvista tekoälytyökaluista, kuten ChatGPT:stä on asiantuntijoille, joilla on tarvittava tieto tunnistaa työkalun generoimat asiavirheet ilman laajempaa faktaselvitystä. Opiskelijoilla ei välttämättä ole tietoa tai osaamista löytää asiantuntevalta kuulostavasta tekstistä asiavirheitä ja lisäksi tekstin todenperäisyyden tarkistamiseen voi mennä jopa enemmän aikaa kuin sen alusta asti itse kirjoittamiseen. Chatbotit voivat parantaa huonon ja keskivertokirjoittajan tekstejä merkittävästi, mutta taitavalle tekstintuottajalle niistä ei kirjoitusmielessä ole suurta hyötyä. Chatbotin tuottama teksti on lisäksi usein generististä ja siitä puuttuu ihmismäinen ote. Pieni joukko meistä tulee oppimaan käyttämään tekoälyä niin hyvin, että saa siitä massiivisen hyödyn, mutta valtaosa meistä tulee jäämään hyvin alkeellisen käyttäjän tasolle, eikä siksi tule saamaan siitä kovinkaan suurta hyötyä. ChatGPT toki pakottaa käyttäjän kehoitteeseen kirjoittaakseen miettimään mikä on ongelma tai kysymys. Jo se itsessään saattaa auttaa hahmottamaan ja purkamaan asiaa pienempiin, ymmärrettäviin osiin.

*pakottaa miettimään
mikä on ongelma tai kysymys*

Esimerkiksi juuri ChatGPT:tä tulisi aineenkirjoituksen sijaan käyttää ajatusten vaihtoon ja perustelemiseen, käsitteiden avaamiseen ja selvittävien lisäkysymysten esittämiseen. Tekoälyä voi käyttää virtuaalitorina,

tuottamaan personoitua oppimateriaalia sekä kokeiden ja luentojen laatimiseen. Oppilaita voi pyytää esimerkiksi kirjoittamaan aineita yhteistyössä ChatGPT:n kanssa ja sen jälkeen pohtimaan, mitä uutta ChatGPT opetti tai missä se teki virheitä.

Jos tekoälyä tarkastellessamme ajattelemme vain ChatGPT:tä ja sen kykyä tehdä oppilaiden esseet ja muut kotitehtävät, ajattelemme hyvin kapea-alaisesti tekoälystä ja sen mukana tuomista mahdollisuuksista ja haasteista. Esimerkiksi ChatGPT:n avulla voi syventää osaamistaan ja parantaa kriittistä ajattelukykyä, ongelmanratkaisutaitoja ja medialukutaitoa, jotka kaikki ovat tärkeitä työelämätaitoja. Teknologian kehittyminen on aiemminkin haastanut tapojamme opettaa. Pitkään koko koulujärjestelmä perustui painetulle kirjalle, ja jossakin määrin perustuu yhä edelleen. Digitalisaation myötä valtaosa kirjoista on korvattu digitaalisella materiaalilla. Tämä ei ole poistanut lukutaidon tarvetta, mutta on helpottanut esimerkiksi lukihäiriöstä kärsivien koulunkäyntiä merkittävästi. Vastaavasti laskimen keksiminen ei poistanut matematiikan osaamisen tarvetta. Kehityksen myötä perustietojen ja -taitojen omaksuminen ennen apuvälineiden käyttöönottoa on noussut avainasemaan. Laskimia ei kouluissa käytetä ennen kuin matematiikan perusosaaminen on tarvittavalla tasolla. Samalla tavalla tietty perusosaamisen taso on hyvä saavuttaa, ennen tekoälytyökalujen käyttöönottoa.

entä esseet?

Uudet teknologiat herättävät keskustelua ja esimerkiksi opetuksen näkökulmasta voidaan pohtia, olisiko aika uudistaa oppimista nykyisestä soveltavampaan suuntaan. On luultavaa, että tekoäly tulee jatkossa olemaan jokaisessa käyttämässämme ohjelmistossa. Tutkimuksen mukaan tekoäly voi vapauttaa 20–40 % peruskouluopettajan ajasta. Mihin ylimenevä aika käytetään? Suurennetaanko luokkakokoja? Jos aika käytetään siihen, että yritämme saada kiinni oppilaita siitä, että he ovat käyttäneet tekoälyä väärin, eikä saavutettu ajallinen hyöty valu hukkaan? Entä jos käyttäisimme ajan oppimista tukeviin toimiin, kuten oppilaiden kanssa vuorovaikutuksessa olemiseen, mentorointiin, ohjaamiseen ja tarjoamaan oppilaille aikaa ja inhimillistä tukea. (Bryant, Heitz, Sanghvi & Wagle; 2020.)

Elon Musk on avoimesti kritisoinut koulutusjärjestelmää. Hänen mielestään tutkinto ei nyky muodossaan kerro siitä, että opiskelija olisi välttämättä oppinut tai että hän olisi poikkeuksellisen lahjakas, vaan siitä, että hän on sinnikäs ja pystyy suoriutumaan kursseista ja kotitehtävistä. ChatGPT on ehkä uusi tapa huijata, mutta huijaaminen itsessään ei ole koulutusmaailmassa uutta. Oppilaat ovat kautta aikojen yrittäneet huijata koetilanteissa ja kotitehtävissä. Jos oppilas on kirjoittanut esseensä naapurillaan, ei opettajilla tähänkään mennessä ole ollut nykyistä kummempia työkaluja oppilaan vilpistä kiinni saamiseen. Miksi oppilaat päätyvät huijaamaan? Ilman motivaatiota oppiminen on paljon haastavampaa. Vailla ymmärrystä siitä, mihin oppimaamme tulemme tarvitsemaan on itsensä motivoiminen äärimmäisen vaikeaa. Tietoa on lähes rajattomasti saatavilla ja oppia voi muullakin tavoin kuin tutkintoja suorittamalla tai esseitä kirjoittamalla. Saavuttaakseen ammattitaidon on kuitenkin osattava tietyt asiat.

“Get good. Stop assigning essays.”

Emad Mostaque, CEO and
Co-Founder of Stability AI

Kun mietitään tekoälytyökalun käyttöönottoa organisaatiossalla, on huomioitava monia asioita. Miten varmistetaan työkalun datan riittävästä laadusta tai siitä, että tietosuojakäytännöt ovat korkealla tasolla? Miten käyttäjien turvallisuus taataan? Esimerkiksi biometrisiä tietoja käytetään tunnistautumiseen monissa sovelluksissa, mutta miten varmistetaan, ettei kenenkään tietoja päädy ei-haluttuihin paikkoihin? Niin opettajien kuin opiskelijoidenkin tulisi sen lisäksi, että oppivat käyttämään tekoälyä, oppia ymmärtämään miten tekoälyn pohjautuvat teknologiat on suunniteltu, miten ne toimivat, sekä ymmärtämään teknologian eettistä kehitystä ja haasteita. Kriittiseksi tarkastelukohteeksi nousee se, miten tämäkin teknologia tulee eriarvoistamaan ihmisiä eri taloudellisilla alueilla, eri talouksissa ja sukupuolienkin välillä. Tällä hetkellä ei ole suunnitelmaa sille, miten tekoälyn, tai ylipäätään digitaalisista työvälineistä saadut hyödyt jakaantuvat kaikkien ihmisten suhteen tasan. Koronapandemia osoitti sen, että digitaalinen jakauma niin kehittyneissä kuin kehittyvissä maissa on suuri, ja valitettavasti pandemia vain pahensi jakaamaa. Tekoälyn maailmassa tarvitaan yhteyksiä, ja ilman niitä digitaalinen jakauma ihmisten välillä tulee kasvamaan entisestään.

digitaalinen eriarvoisuus

Tekoälyä tai teknologiaa ylipäätään ei tulisi nähdä hankalien sosiaalisten ongelmien tai tarpeiden ratkaisijana, siksi on väärin ajatella tekoälyn olevan avain kestävämpään ja oikeudenmukaisempaan koulutukseen. Tästä syystä on erityisen tärkeää miettiä minkälaisia päätöksiä teemme, kun otamme käyttöömme näitä työkaluja. Miten koulutusjärjestelmä muuttuu tekoälyn maailmassa, millainen luokahuoneen tulevaisuus on ja miltä opetus näyttää tulevaisuudessa. Koulutusorganisaatioiden tulee myös ymmärtää, miten niiden sisäiset rakenteet toimivat, kun sen opettajat tai oppilaat käyttävät tällaisia työkaluja. Kuka vastaa siitä, että työkalut ovat linjassa eettisten ja muun koulun ohjeistuksen kanssa? Kuka päättää, mitä nämä ohjeistukset ja linjaukset ovat ja kuka tai mikä on se katto-organisaatio, joka hänen toimintaansa ja päätöksentekoa valvoo? Kuka on vastuussa tietoturvan liittyvistä asioista tai kuka vastaa siitä, että käytettävissä olevat työkalut kohtaavat

Kuka päättää?
Mitä päättää?
Miksi päättää?

ne standardit, jotka niille on asetettu tai kuka organisaatiossa vastaa tekoälytyökaluihin liittyviin kysymyksiin ja huoliin? Kuka on vastuussa opettajien koulutuksesta, että he osaavat käyttää ja hyödyntää työkaluja? Toimivatko työkalut kuten niiden on luvattu tai oletettu toimivan, ja toimivatko ne samalla tavalla kaikille?

Opettajien tulisi harkita, mihin tarkoituksiin he haluavat teknologiaa käyttää ja onko valittu teknologia sopiva käyttötarkoitukseen – mitä käytöllä halutaan saavuttaa oppimisen näkökulmasta ja onko tämän saavuttaminen ylipäätään mahdollista. Tulee myös miettiä, minkä mittaluokan erehdykset ovat hyväksyttäviä. Onko tehtävä sellainen, jossa mahdollinen virhe on sallittu ja mikäli virhe sattuisi, millaiset seuraukset sillä mahdollisesti voisi olla. Kattaako käytön tuoma lisäarvo mahdolliset käytön mukana tuomat riskit? Tärkeää on myös ymmärtää mihin dataan mallit perustuvat ja mikä on työkalun tietolähde. Tulee miettiä, onko data edustavaa ja sopivaa omaan kontekstiin ja käyttötarkoitukseen ja onko datan taustalla esimerkiksi tutkimustietoa. Myös datan vinoumista ja niiden mahdollisesta edistämisestä omalla käytöllä tulee olla tietoinen, ja siitä, kuka omistaa työkalulle syötettävän datan ja mihin tarkoitukseen sitä käytetään.

Tekoälyn ja robotiikan aikakaudella tavat opettaa laajenevat. Opettajien tulee olla hyvin selkeitä oppimistavoitteita määrittellessään. Tekoäly tulee muuttamaan opetustapoja ja opettajien tulee näin ollen pystyä mukauttamaan omaa tapaansa opettaa. Luennot ovat perinteisesti painottuneet passiiviseen oppimiseen, jossa opiskelijat kuuntelevat ja tekevät muistiinpanoja ilman aktiivista ongelmanratkaisua tai kriittistä ajattelua. Tässä muodossaan, pitkien esitysten aikana opiskelijoiden huomio on helposti saattanut herpaantua. Lisäksi lähestymistapa ei huomioi yksilöllisiä eroja ja kykyjä. Tämä on saattanut johtaa joidenkin opiskelijoiden jälkeen jäämiseen opetuksessa, kun taas toiset ovat saattaneet menettää kiinnostuksensa haasteen puutteen vuoksi. Aktiivisessa oppimisessa opiskelijat osallistetaan toiminnallisilla tehtävillä, kuten ongelmanratkaisulla, ryhmätyöskentelyllä ja käytännön harjoituksilla. Tässä lähestymistavassa opiskelijat soveltavat oppimaansa yhteistyössä muiden oppilaiden ja opettajan kanssa. (Mollick, 2023.)

Yksi ratkaisu aktiivisen oppimisen lisäämiseen on flipped learning. Tämä opetusmenetelmä kääntää perinteisen opetusmallin siten, että koulutusmateriaaliin tutustutaan ja harjoitustehtävät tehdään jo ennen varsinaista kontaktiopetusta ja opittua sovelletaan luennolla yhteistyöhön perustuvilla tehtävillä, keskusteluilla tai ongelmanratkaisuharjoituksilla. Flipped learning saattaa toimia hyvin myös tekoälyn yleistyessä. (Mollick, 2023.)

Sen lisäksi että tulee miettiä, miten opetetaan, tulee miettiä mitä opetetaan. On opeteltava opettaa tekoälyä hyödyntäen, ei kieltäen. Jos tietokone voi hoitaa jonkin työn, eikö ole järkevämpää antaa tietokoneen tehdä se? Tämä vapauttaa keskittymään siihen työn osaan, johon tarvitaan ihmisälykkyyttä. Me ihmiset olemme luovia, idearikkaita ja luomme innovaatioita. Jos käytämme työhön apuna tekoälyä, pystymme luultavasti vielä parempiin tuloksiin. Luovuus on ideoiden kantava voima. Luovuutta on vaikeaa opettaa, mutta ideointimenetelmiä opettamalla ruokimme luovuutta. Yhteistyössä tekoälyn kanssa voimme olla entistä luovempia.

Tekoälytutorit voivat räätälöidä opetuksen opiskelijan tarpeisiin ja säätää sisältöä opiskelijan suorituksen perusteella. Kun tekoälytutorit hoitavat osan sisällön jakamisesta, opettajat voivat käyttää enemmän aikaa vuorovaikutuksessa opiskelijoidensa kanssa. Opettajat voivat käyttää tekoälytutoreilta myös tunnistaakseen alueet, joilla opiskelijat tarvitsevat lisätukea tai -ohjausta. Tämä mahdollistaa henkilökohtaisemman ja tehokkaamman opetuksen tarjoamisen. Sitä mukaan, kun tekoälytutorit yleistyvät ja ne yhdistetään flipped learning-malliin voidaan opiskelijoiden oppimiskokemusta merkittävästi parantaa. Tekoäly parantaa viestintää tarjoamalla luonnollisen kielen rajapintoja opiskelijoille kysyä kysymyksiä ja selvennöksiä ilman leimaantumisen pelkoa. Tekoäly voi myös mahdollistaa uusia yhteistyön ja vertaisoppimisen muotoja, yhdistämällä opiskelijoita, joilla on samanlaiset kiinnostuksen kohteet, taustat tai tavoitteet. (Mollick, 2023.)

Tekoälyä apuna käyttäen, opiskelijat voivat säästää aikaa esseiden kirjoittamisessa. Opiskelijat voivat käyttää tuon ajan laajemmin tiedon keräämiseen ja esseen aiheen syvällisempään pohdintaan. Loppujen lopuksi opiskelijoilta voi nyt odottaa enemmän tekoälyn ansiosta. Kaikki eivät välttämättä ole hyviä kirjoittajia. Nyt sitä voi vaatia, koska kaikki voivat tarkistaa tekstinsä kieliopin varalta.

Uusia teknologioita kannattaa vaalia – ne tulevat olemaan osa arkeamme ja ne tulevat kehittymään entisestään. Tämän lisäksi ne tulevat olemaan osa työnteon arkea opiskelijoiden valmistuttua työelämään. Niiden käyttöä kannattaa harjoitella ja niiden heikkouksiin kannattaa tutustua. Opettajien valmiuksien kehittäminen teknologian tehokkaaseen hyödyntämiseen takaa enimmäishyödyn sekä opettajan ajan säästämässä, että oppilaiden tulosten parantamisessa. Opiskelijoiden kanssa voi yhdessä pohtia tekoälytyökalujen mahdollisia käyttökohteita ja millaisia riskejä niiden käyttö voi mahdollisesti tuoda mukanaan. Erilaisten kilpailevien teknologiaratkaisujen moninaisuus tuottaa joskus ongelmia koulu- ja yritysmaailmassa. Siksi on tärkeää jakaa ajatuksia ja kokemuksia siitä mikä toimii ja mikä ei. Työkaluja ja niiden käyttöä kannattaa yhdessä arvioida ja oppia muilta, sekä kollegoilta että opiskelijoilta. Tulevaisuudessa opiskelijoiden, mutta myös opettajien taito olla mukautuva, elinikäinen oppija nousee entistä suurempaan arvoon.

Koska tietoa on saatavilla valtavan paljon, tulee tiedon arvioimisen kyky olemaan entistä tärkeämpi taito tulevaisuudessa. Tekoälynlukutaitoa tulisi opiskelijoille opettaa samalla tavalla kuin medialukutaitoaakin. Mitä pidemmälle generatiivinen tekoäly kehittyy ja mitä enemmän sitä aletaan hyödyntämään, sitä vaikeampaa on erottaa ns. oikeaa tekoälyn tuottamasta. Meidän tulisi olla tietoisia, milloin kommunikoimme tekoälyn kanssa tai mikäli sisältö, jota tarkastelemme, on tekoälyn tuottamaa.

Internetin yleistymisen ja teknologian kehittymisen myötä on ajateltu, että tuottavuutemme nousee uusiin mittasuhteisiin. Vaikka olemme saaneet lisää työkaluja, joiden oletettiin parantavan tuottavuuttamme, ovat ne tuoneet mukanaan myös paljon häiriötekijöitä, eikä tuottavuutemme välttämättä ole merkittävästi ajan myötä kasvanut. Työ on kyllä muuttanut muotoaan ja

helpottunutkin useimmilla aloilla. Sosiaalisten medioiden ajateltiin tuovan ihmisiä yhteen. Olemme yhteydessä toisiimme enemmän kuin koskaan aiemmin, mutta yhteydenpito on hyvin etäistä. Fyysiset ihmiskontaktit vähenivät entisestään pandemian aikana etäopetuksen ja -töiden myötä. Ihmiset ovat yksinäisempiä kuin koskaan ennen.

Koska tekoäly ei nykyisessä muodossaan ole älykäs, se ei pysty esimerkiksi luomaan luottamussuhdetta oppilaaseen ja rakentamaan tätä suhdetta kurssista toiseen. Se ei myöskään pysty ihmisen tavoin ottamaan huomioon mentoroinnissa oppilaan haasteita ja voimavaroja, varsinkaan mikäli niitä ei sille ole opetettu, eikä peilaamaan ajatuksia oppilaslähtöisesti. Monet niistä ominaisuuksista, jotka tekevät hyvistä opettajista loistavia, ovat juuri niitä, joita tekoäly tai muu teknologia ei kykene jäljittelemään, esimerkiksi opiskelijoiden inspiroiminen, positiivisen ilmapiirin luominen, konfliktien ratkaiseminen, yhteenkuuluvuuden tunteen luominen, verkostoituminen sekä opiskelijoiden inhimillinen ohjaaminen. Nämä asiat edustavat opettajan työn ydintä eikä niitä voida eikä pidä automatisoida.

opettajan työn ydintä
ei voi automatisoida

Teknologian kehittyessä opettajien tulee edelleen järjestää oppitunteja joilla oppilaat voivat osallistua keskusteluihin, jakaa mielipiteitä ja oppia toisiltaan. Tekoälyä voidaan käyttää tukena tällaisissa vuorovaikutteisissa toiminnoissa. Opettajan läsnäolo, ohjaus ja vuorovaikutus ovat edelleen välttämättömiä oppilaiden kokonaisvaltaisen kehityksen kannalta. Inhimillinen vuorovaikutus edistää oppilaiden sosiaalisia taitoja, empatiaa ja ryhmätyövalmiuksia – tärkeitä taitoja työelämää varten.

Perusmallit, kuten OpenAI:n GPT, Googlen PaLM ja Metan LLaMa ovat yleiskäyttöisiä laskentalaitteita. Ne voivat vastata eri asteisella menestyksellä lähes mihin tahansa kysymykseen, jonka niille esittää. Nykyiset mallit koulutusmenetelmien ovat melko yksinkertaisia. Perusmallien kehityksessä, tutkijoiden oletetaan löytävän lisää tapoja hyödyntää niitä datan, mallien ja koulutusprosessien parantamiseen. On paljon varsin ilmeisiäkin tapoja parantaa perusmalleja merkittävästi, ja siksi ei ole syytä uskoa, että olemme lähelläkään teknologian rajoja. Teknologian rajoja tai sitä, kuinka vaikeaa tai hidasta niiden saavuttaminen tulee olemaan, on hankala ennustaa. Emme voi tehdä ennusteita siitä, kuinka pitkälle tämä kehitys voi edetä tai mitä tulevaisuudessa tapahtuu. Monet alan tutkijat ja johtajat uskovat, että käytännön rajoituksia ei välttämättä ole ollenkaan. (Howard, 2023.) On vain todettava, ettei meillä oikeastaan ole mitään käsitystä siitä mihin tämä vielä voi johtaa.

LÄHTEET

Arene, 2023. Arenen suosituksiset tekoölyn hyödyntämisestä ammattikorkeakouluille. Saatavissa: <https://arene.fi/wp-content/uploads/PDF/2023/AI-Arene-suositukset.pdf?t=1686309593> Viitattu 18.9.2023.

Bryant, J; Heitz, C; Sanghvi, S & Wagle, D. 2020. How artificial intelligence will impact K–12 teachers. McKinsey Insights. Saatavissa: <https://www.mckinsey.com/industries/education/our-insights/how-artificial-intelligence-will-impact-k-12-teachers>. Viitattu 20.6.2023.

Choi, Y. 2023. Why AI is incredibly smart and shockingly stupid. Ted Talks. Saatavissa: https://www.ted.com/talks/yejin_choi_why_ai_is_incredibly_smart_and_shockingly_stupid/c. Viitattu 3.7.2023.

Chui, M; Roberts, R; Rodchenko, T; Singla, A; Sukharevsky, A; Yee, L & Zurkiya, D. 2023. What every CEO should know about generative AI. McKinsey & Company. Saatavissa: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/what-every-ceo-should-know-about-generative-ai/#/>. Viitattu 27.7.2023.

Commoditizing the Petaflop - with George Hotz of the tiny corp (Latent Space). 2023. Saatavissa: <https://www.latent.space/p/geohot#details>. Viitattu 11.7.2023.

Dastin, J. & Nellis, S. 2023. Focus: For tech giants, AI like Bing and Bard poses billion-dollar search problem. Reuters. Saatavissa: <https://www.reuters.com/technology/tech-giants-ai-like-bing-bard-poses-billion-dollar-search-problem-2023-02-22/>. Viitattu 1.8.2023.

Davis, W. 2023. Google's medical AI chatbot is already being tested in hospitals. The Verge. Saatavissa: <https://www.theverge.com/2023/7/8/23788265/google-med-palm-2-mayo-clinic-chatbot-bard-chatgpt>. Viitattu 12.7.2023.

Dickson, B. 2023. How open-source LLMs are challenging OpenAI, Google, and Microsoft. TechTalks. Saatavissa: <https://bdtechtalks.com/2023/05/08/open-source-llms-moats/>. Viitattu 12.7.2023.

Digivisio 2030. 2022. Neljä väitettä oppimisen tulevaisuudesta. Millaisen maailman oppimisen asiantuntijat näkevät 2030-luvulla. Digivisio 2030. Saatavissa: https://www.youtube.com/watch?v=J_F11JlvK70. Viitattu 30.5.2023.

Dunn, C. 2023. OpenAI tokens and limits. Saatavissa: <https://devblogs.microsoft.com/surface-duo/android-openai-chatgpt-15/>. Viitattu 11.9.2023.

Euroopan parlamentti. 2023. EU-parlamentti hyväksyi kantansa: tekoölyn käytön oltava turvallista ja avointa. Lehdistötiedote. Saatavissa: <https://www.europarl.europa.eu/news/fi/press-room/20230609IPR96212/eu-parlamentti-tekoalyn-kayton-oltava-turvallista-ja-avointa>. Viitattu 29.6.2023.

European parliament. 2023. Artificial intelligence act. Saatavissa: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf). Viitattu: 4.7.2023.

Elements of AI. 2018a. Mitä tekoöly on? MinnaLearn ja Helsingin yliopisto. Verkkokurssi. Saatavissa: <https://course.elementsofai.com/fi/1/3>. Viitattu 1.6.2023.

Elements of AI. 2018b. Ongelmanratkaisu tekoälyn avulla. MinnaLearn ja Helsingin yliopisto. Verkkokurssi. Saatavissa: <https://course.elementsofai.com/fi/2>. Viitattu 6.6.2023.

Elements of AI. 2018c. Koneoppiminen. MinnaLearn ja Helsingin yliopisto. Verkkokurssi. Saatavissa: <https://course.elementsofai.com/fi/4>. Viitattu 6.6.2023.

Elements of AI. 2018d. Tekoälyn käytännön sovellukset. MinnaLearn ja Helsingin yliopisto. Verkkokurssi. Saatavissa: <https://course.elementsofai.com/fi/4>. Viitattu 7.6.2023.

Elements of AI. 2018e. Neuroverkot. MinnaLearn ja Helsingin yliopisto. Verkkokurssi. Saatavissa: <https://course.elementsofai.com/fi/4>. Viitattu 8.6.2023.

Elements of AI. 2018f. Tekoälyn vaikutukset. MinnaLearn ja Helsingin yliopisto. Verkkokurssi. Saatavissa: <https://course.elementsofai.com/fi/4>. Viitattu 8.6.2023.

FCAI. 2023. EU:n lakiesitys tekoälyjärjestelmien säätelemisestä sisältää vakavia puutteita. Finnish Center for Artificial Intelligence. Saatavissa: <https://fcai.fi/news-in-finnish/2021/4/15/eun-lakiesitys-tekoalyjarjestelmien-saatelemisesta>. Viitattu 5.6.2023.

Grady, P & Castro, D. 2023. Tech Panics, Generative AI, and the Need for Regulatory Caution. Saatavissa: <https://datainnovation.org/2023/05/tech-panics-generative-ai-and-regulatory-caution/#fn88>. Viitattu 22.8.2023.

Hotakainen, O & Kivinen, J. 2023. Tekoälypohjainen oppijan ohjaus ja neuvonta. Digivisio 2030. Saatavissa: https://digivisio2030.fi/wp-content/uploads/2023/06/Tekoalypohjainen_oppijan_ohjaus_ja_neuvonta-lopullinen-1.pdf. Viitattu 7.6.2023.

Hoffmann, J; Borgeaud, J; Mensch, A; Buchatskaya, E; Cai, T; Rutherford, E; de Las Casas, D; Hendricks, L A; Welbl, J; Clark, A; Hennigan, T; Noland, E; Millican, K; van den Driessche, G; Damoc, B; Guy, A; Osindero, S; Simonyan, K; Elsen, E; Rae, J.W; Vinyals, O & Sifre, L. 2022. Training Compute-Optimal Large Language Models. DeepMind. Saatavissa: <https://arxiv.org/pdf/2203.15556.pdf>. Viitattu: 11.7.2023.

Howard, J. 2023. AI Safety and the Age of Dislightenment. Saatavissa: <https://www.fast.ai/posts/2023-11-07-dislightenment.html>. Viitattu 1.8.2023.

Humanity is not that simple. 2023. Yuval Noah Harari & Pedro Pinto. Saatavissa: <https://www.youtube.com/watch?v=4hllDiVDww4>. Viitattu 28.6.2023.

Hursti, M; Puuppo, J; Raita-Hakola, A-M; Riihiaho, K; Lind, L; Hämäläinen, J; Kettunen, L; Uotinen, V; Waselius, T & Pölönen, I. 2023a. Tekoäly opetuksen tukena-verkkokurssi. Tekoälyn perusteet. Jyväskylän yliopisto. Saatavissa: <https://tim.jyu.fi/view/kurssit/tie/ai-ja-chatgtp/tekoalyn-perusteet>. Viitattu 15.6.2023.

Hursti, M; Puuppo, J; Raita-Hakola, A-M; Riihiaho, K; Lind, L; Hämäläinen, J; Kettunen, L; Uotinen, V; Waselius, T & Pölönen, I. 2023b. Tekoäly opetuksen tukena-verkkokurssi. Tekoälyn opettajan arjessa. Jyväskylän yliopisto. Verkkokurssi. Saatavissa: <https://tim.jyu.fi/view/kurssit/tie/ai-ja-chatgtp/chat-gpt-opettajan-arjessa>. Viitattu 15.6.2023.

Jarvis, R. 2023. OpenAI CEO, CTO on risks and how AI will reshape society. ABC News Exclusive. Saatavissa: <https://abcnews.go.com/Technology/video/openai-ceo-cto-risks-ai-reshape-society-97949497>. Viitattu 30.5.2023.

Jones, R. 2023. Nick Clegg Is Driving Meta's Case for Open-Sourced AI. Observer. Saatavissa: https://observer.com/2023/07/nick-clegg-meta-open-sourced-ai/?utm_source=tldr.ai. Viitattu 13.7.2023.

Kalliamvakou, E. 2022. Research: quantifying GitHub Copilot's impact on developer productivity and happiness. GitHub. Saatavissa: <https://github.blog/2022-09-07-research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/>. Viitattu 31.5.2023.

Kharpal, A. 2023a. China's Baidu claims its Ernie Bot beats ChatGPT on key tests as A.I. race heats up. CNBC Tech. Saatavissa: <https://www.cnbc.com/2023/06/27/baidu-claims-its-ernie-bot-beats-openais-chatgpt-on-key-ai-tests.html>. Viitattu 8.8.2023.

Kharpal, A. 2023b. Chinese tech giant Alibaba challenges Meta with open-sourced A.I. model launch. CNBC Tech. Saatavissa: https://www.cnbc.com/2023/08/03/alibaba-launches-open-sourced-ai-model-in-challenge-to-meta.html?utm_source=tldr.ai. Viitattu 8.8.2023.

Kleczek, D. 2023. A Gentle Introduction to LLM APIs. Saatavissa: <https://wandb.ai/darek/llmapps/reports/A-Gentle-Introduction-to-LLM-APIs--VmlldzoONjM0MTMz>. Viitattu 11.9.2023.

Knight, W. 2023. OpenAI's CEO Says the Age of Giant AI Models Is Already Over. Saatavissa: <https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/>. Viitattu 11.7.2023.

Lawton, G. 2023. Generative models: VAEs, GANs, diffusion, transformers, NeRFs. TechTarget. Saatavissa: <https://www.techtarget.com/searchenterpriseai/tip/Generative-models-VAEs-GANs-diffusion-transformers-NeRFs>. Viitattu: 6.7.2023.

Lukianov, D. 2023. A Closer Look at Large Language Models. Saatavissa: <https://akvelon.com/a-closer-look-at-large-language-models/>. Viitattu 8.9.2023.

Merilampi. 2023a. Tekoälysäätely on paljon muutakin kuin EU:ssa valmisteilla oleva tekoälyasetus. Saatavissa: <https://www.merilampi.com/tekoalysaantely-on-paljon-muutakin-kuin-eussa-valmisteilla-oleva-tekoalysasetus/?lang=fi>. Viitattu 30.5.2023.

Merilampi. 2023b. Miten pitäisi valmistautua EU:n tekoälysäätelyyn? Saatavissa: <https://www.merilampi.com/miten-pitaisi-valmistautua-eun-tekoalysaantelyyn/?lang=fi>. Viitattu 5.6.2023.

Meta AI. 2023. Introducing Llama 2. Saatavissa: https://ai.meta.com/llama/?utm_source=tldr.ai. Viitattu 20.7.2023.

Microsoft, 2023. 2022 Environmental Sustainability Report. Saatavissa: <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW15mgm> Viitattu 13.9.2023.

Mollick, E. 2023. The future of education in a world of AI. Saatavissa: <https://www.oneusefulthing.org/p/the-future-of-education-in-a-world> Viitattu 19.9.2023.

Muehmel, 2023. What Is a Large Language Model, the Tech Behind ChatGPT? Saatavissa: <https://blog.dataiku.com/large-language-model-chatgpt> Viitattu 5.9.2023.

Murto, M. 2023. Henkilökohtainen tiedonanto, Teams-keskustelu. 19.9.2023.

Myers West, S & Kak, A. 2023. 2023 Landscape – Confronting Tech Power. AI Now Institute. Saatavissa: <https://ainowinstitute.org/wp-content/uploads/2023/04/AI-Now-2023-Landscape-Report-FINAL.pdf>. Viitattu 28.6.2023.

Myers West, S & Vipra, J. 2023. Computational Power and AI. AI Now Institute. Saatavissa: <https://ainowinstitute.org/publication/policy/computational-power-and-ai>. Viitattu 26.6.2023.

Nomic. 2023. GPT4All A free-to-use, locally running, privacy-aware chatbot. Saatavissa: <https://gpt4all.io/index.html>. Viitattu 12.7.2023.

O'Brien, M & Fingerhut, H, 2023. The Associated Press. Artificial intelligence technology behind ChatGPT was built in Iowa — with a lot of water. Saatavissa: <https://apnews.com/article/chatgpt-gpt4-iowa-ai-water-consumption-microsoft-f551fde98083d17a7e8d904f8be822c4> . Viitattu 13.9.2023.

Open AI. 2023. Tokenizer. Saatavissa: <https://platform.openai.com/tokenizer>. Viitattu 15.9.2023.

Palmer, W, Walsh, B. & Nair, N. 2023. The Impact of Generative AI and Large Language Models on Organizational Sustainability and ESG Goals. World Wide Technology. Saatavissa: <https://www.wwt.com/article/the-impact-of-generative-ai-and-large-language-models-on-organizational-sustainability-and-esg-goals>. Viitattu 25.7.2023.

Patel, D & Ahmad, A. 2023a. The Inference Cost Of Search Disruption – Large Language Model Cost Analysis. SemiAnalysis. Saatavissa: <https://www.semianalysis.com/p/the-inference-cost-of-search-disruption>. Viitattu 28.6.2023.

Patel, D & Ahmad, A. 2023b. Google “We Have No Moat, And Neither Does OpenAI”. SemiAnalysis. Saatavissa: <https://www.semianalysis.com/p/google-we-have-no-moat-and-neither>. Viitattu 12.7.2023.

Peters, J & Porter, J. 2023. Instagram’s Threads surpasses 100 million users. The Verge. Saatavissa: <https://www.theverge.com/2023/7/10/23787453/meta-instagram-threads-100-million-users-milestone>. Viitattu 10.7.2023.

Rakennerahastotietopalvelu. Euroopan sosiaalirahaston (ESR) rahoittaman hankkeen kuvaus. Saatavissa: <https://www.eura2014.fi/rrtiepa/projekti.php?projektikoodi=S22683> Viitattu 18.9.2023.

Schreiner, M. 2023. GPT-4 architecture, datasets, costs and more leaked. The decoder. Saatavissa: <https://the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/>. Viitattu 19.7.2023.

Shinde, S. 2023. What Companies Fall Under Big Tech? How Do You Land a Job With Them? Saatavissa: <https://emeritus.org/blog/technology-big-tech/>. Viitattu 20.7.2023.

Shrink that footprint. 2022. Carbon Footprint Of Training GPT-3 And Large Language Models. Saatavissa: <https://shrinkthatfootprint.com/carbon-footprint-of-training-gpt-3-and-large-language-models/>. Viitattu: 25.7.2023.

Silberg, J & Manyika, J. 2019. Tackling bias in artificial intelligence (and in humans). McKinsey Global Institute. Saatavissa: <https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans>. Viitattu 26.6.2023

Solita. 2023. Solita summit 2023. Lead in the new era of AI – key trends and considerations for business to succeed. Saatavissa <https://youtu.be/fBvLSFHGso4>. Viitattu 31.5.2023.

Tam, A. 2023. What are Large Language Models. Guiding Tech Media. Machine learning mastery. Saatavissa: <https://machinelearningmastery.com/what-are-large-language-models/>. Viitattu 26.6.2023.

Thomas, A. 2023. Data and Compute Are the Ultimate Flywheel. Saatavissa: <https://every.to/p/data-in-the-age-of-ai>. Viitattu 7.9.2023.

Toews, R. 2023. Transformers Revolutionized AI. What Will Replace Them? Saatavissa: <https://www.forbes.com/sites/robtoews/2023/09/03/transformers-revolutionized-ai-what-will-replace-them/> Viitattu 5.9.2023.

Toner, H. 2023. What are generative AI, large language models, and foundation models? Center for security and emerging technology. Saatavissa: <https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/>. Viitattu 7.9.2023.

Touvron, H; Lavril, T; Izacard, G; Martinet, X; Lachaux, M-A; Lacroix, T; Rozière, B; Goyal, N; Hambro, E; Azhar, F; Rodriguez, A; Joulin, A; Grave, E; Lample, G. 2023. Meta AI. LLaMA: Open and Efficient Foundation Language Models. Saatavissa: <https://arxiv.org/pdf/2302.13971.pdf>. Viitattu 15.9.2023.

TurkuNLP. 2023. FinGPT-3. Saatavissa: <https://turkunlp.org/gpt3-finnish>. Viitattu 8.8.2023.

Vanian, J. 2023. Microsoft goes beyond OpenAI, makes Meta's new A.I. model available to Azure customers. CNBC Tech. Saatavissa: <https://www.cnbc.com/2023/07/18/microsoft-makes-metas-new-ai-model-available-to-azure-customers.html>. Viitattu 8.8.2023.

Verma, P. 2023. What to know about OpenAI, the company behind ChatGPT. The Washington Post. Saatavissa: <https://www.washingtonpost.com/technology/2023/02/06/what-is-openai-chatgpt/>. Viitattu: 19.7.2023.

What are Generative AI models? (IBM Technology) 2023. Saatavissa: <https://www.youtube.com/watch?v=hflUstzHs9A&t=499s>. Viitattu 7.9.2023.

Why AI Matters And How To Deal With The Coming Change w/ Emad Mostaque. (Peter H. Diamandis) 2023. YouTube. Saatavissa: https://www.youtube.com/watch?v=ciX_iFGyS0M. Viitattu 12.7.2023.

World Economic Froum. 2023a.The Golden age of AI: why ChatGPT is just the start. Radio Davos. Podcast. Saatavissa: https://open.spotify.com/episode/7uQGHvRBjzZ3hP6NFRCSr?si=Pdygx_wOSWmnjE_CT-jh77Q50446799 . Viitattu 1.6.2023.

World Economic Froum. 2023b. Rise of the machines: Prof. Stuart Russell on the promises and perils of AI. Radio Davos. Podcast. Saatavissa: <https://open.spotify.com/episode/0VLO9mldvzsWu9MYEYbujc?si=conz1g-trSFG9jg4gJ3fiZQ>. Viitattu 1.6.2023.

World Economic Froum. 2023c. Growth Summit 2023 - Generation AI – English. Saatavissa: <https://www.weforum.org/videos/growth-summit-2023-generation-ai-english>. Viitattu 2.6.2023.

World Economic Froum. 2023d. The rise of AI and the green transition will transform the way we work: Future of Jobs Report 2023. Radio Davos. Podcast. Saatavissa https://open.spotify.com/episode/1o-lmMsHRPY-8f1ToT29e2u?go=1&sp_cid=5d9e2225a221c6888c4752b7c98403cf&t=12&utm_source=embed_player_p&utm_medium=desktop&nd=1. Viitattu 2.6.2023.

Yuval Noah Harari, 2022. Frequently asked questions - answered by Yuval Noah Harari. Yahav-Harari Group Ltd. Saatavissa: <https://www.ynharari.com/faqs/>. Viitattu 28.7.2023.

Zola, A & Vaughan, J. 2022. Definition: backpropagation algorithm. TechTarget. Saatavissa: <https://www.techtarget.com/searchenterpriseai/definition/backpropagation-algorithm>. Viitattu 12.6.2023.

centria
ammattikorkeakoulu

modo
Modulaarinen, digitaalinen
ja alykäs oppimisympäristö



Euroopan unioni

Euroopan sosiaalirahasto

Vipuvoimaa
EU:lta
2014–2020



Elinkeino-, liikenne- ja
ympäristökeskus

Hanketta rahoitetaan Euroopan sosiaalirahastosta osana Euroopan unionin COVID-19 -pandemian johdosta toteuttamia toimia.