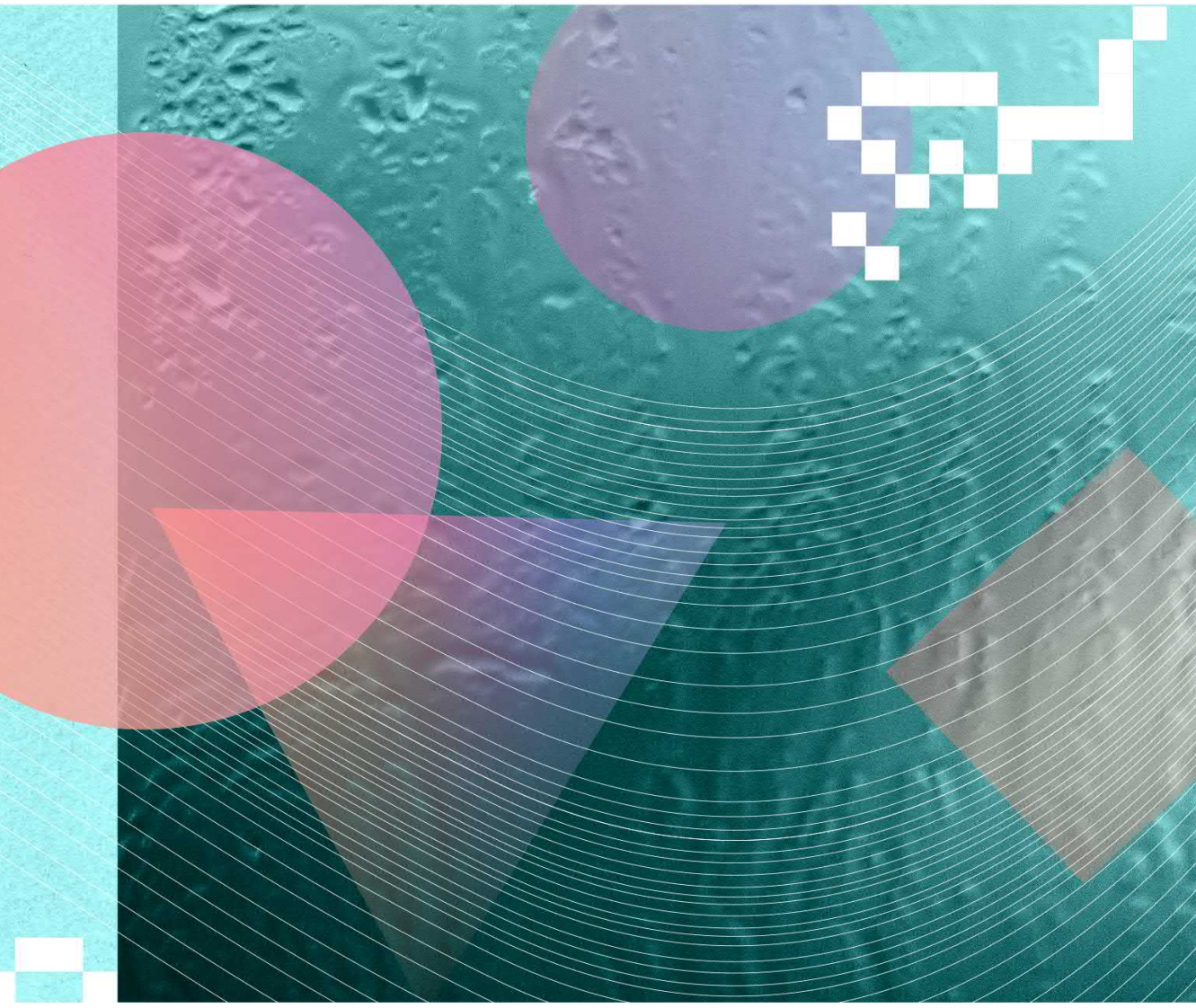
EDUSKUNNAN TULEVAISUUSVALIOKUNNAN JULKAISU 1/2022

TEKOÄLYT TÄNÄÄN JA TULEVAISUUDESSA (työnimi)



Tulevaisuusvaliokunta

00102 Eduskunta
www.eduskunta.fi

Helsinki

ISBN 978-951-53-3781-8 (nid.)
ISBN 978-951-53-3782-5 (pdf)

ISSN 2342-6594 (painettu)
ISSN 2342-6608 (verkkojulkaisu)

Tekoälyt tänään ja tulevaisuudessa (työnimi)

Heikki Ailisto & Petri Myllymäki,
Sasu Tarkoma, Joni-Kristian Kämäräinen,
Juha Röning, Tapio Salakoski, Arno Solin,
Pertti Saariluoma,Tommi Mikkonen,
Mark van Gils, Kaisa Väänänen, Kai Puolamäki,
Peter Ylén, Teemu Roos, Jaana Leikas,
Antti Honkela, Matti Kutila, Laura Ruotsalainen,
Petri Ylikoski, Risto Linturi

EDUSKUNNAN TULEVAISUUSVALIOKUNNAN JULKAISU 1/2022

Sisällys

TULEVAISUUSVALIOKUNNAN ESIPUHE 8

OSA I TEKOÄLYTUTKIJOIDEN NÄKÖKULMIA 9

1 **Johdanto: Mitä tekoäly on** | Heikki Ailisto & Petri Myllymäki 10
1.1 Mitä tekoäly on 10
1.1.1 Koneoppiminen neuroverkoilla 12
1.1.2 Koneoppimisen haasteet 13
1.2 Tekoäly muutosvoimana 14
1.3 Viestit päättäjille ja kansalaisille 16

2 **Tekoäly tehostaa ja analysoi** | Sasu Tarkoma 17
2.1 Mistä on kyse? 17
2.2 Tekoäly muutosvoimana 19
2.3 Viestit päättäjille ja kansalaisille 20

3 **Tekoäly havainnoi – koneaistit** | Joni-Kristian Kämäräinen 23
3.1 Tekoäly havainnoi – tulin, näin, voitinko? 23
3.2 Koneaistit – tietokonenäkö tekoälyn veturina 23
3.2.1 Yksi menetelmä löytää heidät, se yksi heitä hallitsee 25
3.3 Tekoälyn nykytila – kun Suomi putos puusta 25
3.3.1 Tutkimus siirtynyt yrityksiin – kaiken takana on data, joka ihanasti opettaa 26
3.4 Tehdään Suomesta taas suuri – oppia elokapinalta ja Gretzkyltä 27
3.4.1 Eurooppalainen vastarintaliike, datakapina 27
3.4.2 Pienen Suomen pitää tehdä sitä, missä se on hyvä, eikä sitä, mikä on kivaa 28
3.4.3 Mihin kiekko menee seuraavaksi? Luonnollinen kieli ja robotiikka 28

4 **Tekoäly liikkuu ja tekee** | Juha Röning 30
4.1 Tekoäly muutosvoimana: Tekoälyn, datan ja robotiikan mahdollisuudet 31
4.1.1 Robotiikan tutkimus Euroopassa 31
4.1.2 Tekoäly – hyvä renki, mutta huono isäntä 32
4.1.3 Tekoälyn mahdollisuudet robotiikassa 33
4.2 Viestit päättäjille ja kansalaisille 35

5 Tekoäly lukee, kirjoittaa ja keskustelee | Tapio Salakoski 36
5.1 Mistä on kyse? 36
5.1.1 Puhe ja teksti 36
5.1.2 Kielii ja kielimallit 37
5.2 Kieliteknologian käyttötapoja 37
5.2.1 Tiedonhaku 37
5.2.2 Sisältöanalyysi 38
5.2.3 Tekstin tuottaminen ja vaikuttaminen 38
5.3 Tekoäly muutosvoimana 39
5.3.1 Pienten kielten asema 39
5.3.2 Digitalisaation tehostuminen 40
5.3.3 Yksityisyys, väärinkäytökset, valvonta ja vaikuttaminen 40
5.4 Viestit päättäjille ja kansalaisille 40
5.4.1 Huippuosaaminen ja yleissivistys 40
5.4.2 Sääntely ja kilpailukyky 40
5.4.3 Kokonaisturvallisuus 41
5.5 Automatisoituvat kielipalvelut 41
5.6 Kilpailukykyä kieliälyllä 42

6 Tekoäly oppii | Arno Solin 43
6.1 Mistä on kyse? 43
6.1.1 Ohjaamaton oppiminen 44
6.1.2 Ohjattu oppiminen 45
6.1.3 Vahvistusoppiminen 45
6.2 Tekoäly muutosvoimana 46
6.3 Huomioitavaa päättäjille ja kansalaisille 46
6.4 Oppivat tekoälyt täydentävät ihmisen osaamista 47

7 Tekoäly luo | Pertti Saariluoma 48
7.1 Johdanto 48
7.2 Miksi tekoälyjärjestelmät voivat tuottaa luovia tuloksia? 49
7.3 Tekoäly avustaa luovia toimintoja 50
7.4 Käytännön esimerkkejä 51
7.5 Luova uudistaminen tulevaisuudessa 52

Sisällys

8 Tekoäly kansalaisten palveluksessa | Tommi Mikkonen 54
8.1 Johdanto 54
8.2 Tilannekuva 54
8.2.1 Kansalaisten tarpeet 54
8.2.2 Kansalaisille tarjotut tietojärjestelmät ja niiden data 55
8.3 Vaihtoehtoiset etenemistavat 56
8.3.1 Vaihtoehto 1: Paikalliset pilotit 56
8.3.2 Vaihtoehto 2: Teknologia- ja toteutusvetoiset toimenpiteet 57
8.3.3 Vaihtoehto 3: Strategiset toimenpiteet 58
8.4 Yhteenveto ja suosituksia 58

9 Tekoäly terveydenhuollossa | Mark van Gils 60
9.1. Mistä on kyse? 60
9.2 Tekoäly muutosvoimana 63
9.3 Viestit päättäjille ja kansalaisille, suositukset 64
9.4 Terveydenhuolto ja tekoäly – kolme haastetta, joihin tarvitaan apua 65

10 Tekoäly kaupunkikehityksessä
– ihmiskeskeisen suunnittelun näkökulma | Kaisa Väänänen 66
10.1 Mistä ihmiskeskeisessä tekoälysuunnittelussa on kyse 67
10.2 Tekoäly kestävän kaupunkikehityksen muutosvoimana 68
10.3 Viestit päättäjille ja kansalaisille:
Ihmiskeskeisyys on avaintekijä tekoälyn vaikuttavuuteen 70
10.3.1 Politiikkasuosituksia päättäjille 70
10.3.2 Suosituksia kansalaisille 70
10.4 Botit auttavat ja kannustavat kaupunkilaisia 71

11 Tekoäly tieteellisessä tutkimuksessa | Kai Puolamäki 73
11.1 Mistä on kyse? 73
11.2 Tekoäly muutosvoimana 75
11.3 Viestit päättäjille ja kansalaisille 76

12 Tekoäly syrjäytymisen estämisessä | Peter Ylén 77
12.1 Mistä on kyse? 77
12.2 Tekoäly muutosvoimana. 78
12.2.1 Syrjäytymisen data ja etiikka 78
12.2.2 Virtuaaliset tietoaltaat 80
12.2.3 Tekoälyn mahdollisuudet 80
12.3 Toimenpide-ehdotukset 81

13 Tekoälystä ymmärrettävää | Teemu Roos 82
13.1 Tekoälyn lukutaito eli mitä jokaisen tulisi tietää tekoälystä 82
13.2 Tekoälyn poliittiset ulottuvuudet 84
13.3 Tekoälyn hyödyntäminen poliittisessa päätöksenteossa 85
13.4 Yhteenveto ja politiikkasuosituksia 85
13.5 Miksi tekoäly ei ole asiantuntija 86

14 Tekoälystä vastuullista | Jaana Leikas 88
14.1 Mistä on kyse? 88
14.1.1 Tekoälyn eettiset kysymykset ovat moniuloitteisia 88
14.1.2 Eettiset koodistot ja ohjeet viitoittavat suuntaa 89
14.2 Tekoäly muutosvoimana 90
14.2.1 Tekoäly haastaa luottamusta 90
14.2.2 Lisää näkemystä kokonaisuuden hallintaan 90
14.2.3 Lainsäädännön rinnallle pehmeää ohjausta 91
14.2.4 Suunnittelijan vastuu on suuri 92
14.3 Viestit päättäjille ja kansalaisille 92
14.4 Tekoäly-yhteiskunnan etiikkaa 94

15 Tekoälystä tietoturvallista | Antti Honkela 95
15.1 Mistä on kyse 95
15.1.1 Tekoäly ja tietosuoja sekä yksityisyyden suoja 96
15.1.2 Tekoälyn tietoturva 97
15.1.3 Ihminen tekoälyn valvojana ja yhteistyökumppanina 98
15.2 Tekoäly muutosvoimana 98
15.3 Viestit päättäjille ja kansalaisille 99
15.4 Kansalaisten tietosuojan turvaava tietojen jako 101

Sisällys

16 Tekoälystä fyysisesti turvallista: esimerkkinä automaattiset ajoneuvot | Matti Kutila 102
16.1 Mistä on kyse? 102
16.2 Tekoäly muutosvoimana 105
16.3 Toimenpiteet tekoäly -pohjaisen automaation edistämiseen 108
16.4 Automaattiset ajoneuvot 110

17 Tekoäly kestävän kehityksen ympäristönäkökulmasta | Laura Ruotsalainen 111
17.1 Kestävän kehityksen tavoitteet ja haasteet 111
17.1.1 Tekoäly edistämässä kestävän kehityksen ympäristötavoitteita 112
17.1.2 Älykaupungit ja automaattiliikenne esimerkkeinä kestävyysmurroksesta 113
17.2 Tekoälyyn perustuvien ratkaisujen haasteet kestävän kehityksen näkökulmasta 114
17.3 Toimenpide-ehdotukset 115
17.4 Case-esimerkki: Tekoälyn avulla kohti kestävyysmurrosta ilmastonmuutoksen torjumiseksi ja luonnon monimuotoisuuden säilyttämiseksi 116

18 Yhteiskunnallisesti kestävä tekoäly | Petri Ylikoski 117
18.1 Mistä on kyse? 117
18.2 Kestävän tekoälyn edellytys 118
18.2.1 Luotettava data 118
18.3 Viestit päättäjille ja kansalaisille 120
18.3.1 Politiikka suositukset 121

II TULEVAISUUSVALIOKUNNAN KOKEILU TEKSTIÄ TUOTTAVASTA JA DIALOGISESTA TEKOÄLYSTÄ TUKIÄLYN NÄKÖKULMASTA 122

19 Johdanto tulevaisuusvaliokunnan kokeiluun tekstiä tuottavasta ja dialogisesta tekoälystä tukiälyn näkökulmasta | Risto Linturi 123
19.1 Mistä on kyse? 123
19.2 Tekoäly muutosvoimana – kaksi polkua tulevaisuuteen 125
19.3 Viestit päättäjille ja kansalaisille, valitkaa viisaasti: 127

20 Raportti tulevaisuusvaliokunnan kokeilusta: tekstiä tuottava ja dialoginen tekoäly (lausunto 11.4.2021) | Risto Linturi 128
20.1 Tulevaisuusvaliokunnan kokeilu tekstiä tuottava ja dialoginen tekoäly 128
20.1.1 Tausta 128
20.1.2 Sisältö 128
20.2 Tekoälyn lyhyt historia ja tärkeimmät periaatteet 129
20.3 Tekoälyn nykyiset merkittävät sovellusalueet esimerkinomaisesti 131
20.4 Tekstiä generoivat järjestelmät, erityisesti GPT-3 tekstin tuotannossa 132
20.5 Tekstin jatkamisesta ihmisen ja koneen vuoropuheluun 134
20.6 Tekoälyasiantuntijakokeilun kolme eri vaihetta 134
20.6.1 Edustajien etukäteen lähettämät kysymykset 137
20.6.2 Keskustelut valiokunnan kuulemistilaisuudessa 139
20.7 Näkemys GPT-3:n ja Project December -tyyppisten persoonallisuusmatriisien hyödyt 140
20.8 Tekoälyn kehitysnäkymiä, uhkia ja mahdollisuuksia 140

LIITTEET
Liite 1 Muskien ja Saaran kanssa käydyt valmistelevat harjoituskeskustelut 143
Liite 2 Muskien ja Saaran kanssa käydyt keskustelut edustajien etukäteen lähettämistä kysymyksistä 148
Liite 3 Valiokunnan kokouksessa käydyt keskustelut 155

Valiokunnan alkusanat

Tähän sitten niin monta sivua, kun tarvitaan, huom, lopuksi sivunumerot!

OSA I
Tekoälytutkijoiden näkökulmia

1 Johdanto: Mitä tekoäly on?

Heikki Ailisto
tutkimusprofessori, Teknologian tutkimuskeskus VTT Oy
Tekoälyn tutkimuskeskus FCAI, vastuualueena teollisuus ja yhteiskunta

Petri Myllymäki
professori (tekoäly ja koneoppiminen), Helsingin yliopisto
Varajohtaja, tekoälyn tutkimuskeskus FCAI

Tässä johdantoluvussa kuvataan tekoälyä yleisellä tasolla. Siinä haetaan vastausta kysymykseen, mikä on tekoälyn ja tavallisen tietokoneohjelman ero, ja käsitellään tekoälyä muutosvoimana taloudessa ja laajemmin yhteiskunnassa. Johdantoluvussa tuodaan myös esille seikkoja, jotka sekä päättäjien että kansalaisten on tekoälyn ja digitalisaation ajamassa murroksessa syytä ottaa huomioon.

1.1 Mitä tekoäly on?

Tekoäly nähdään yhtenä merkittävimmistä teknologisista muutosvoimista. Siihen liitetään yhtäältä suuria toiveita esimerkiksi talouden kasvusta, ympäristön muutosten hallinnasta sekä lääketieteen ja muiden alojen tieteellisistä läpimurroista mutta toisaalta uhkakuvia massatyöttömyydestä, eriarvoistumisesta, valvontayhteiskunnasta ja jopa perustavaa laatua olevasta uhasta ihmiskunnalle. On siis syytä yrittää selvittää, mitä tekoäly on.

Tekoälyn historia tieteenalana ulottuu 1950-luvulle. Tutkijat asettivat tavoitteeksi inhimillisen kaltaiseen päättelyyn, järkeilyyn ja toimintaan pystyvien keinotekoisten (*artificial*) järjestelmien rakentamisen. Vasta kehitettyjen tietokoneiden yhdistämisen kehittyviin kognitiotieteisiin, perinteiseen logiikkaan, lingvistiikkaan ja muihin tutkimusaloihin arveltiin tuottavan keinotekoista älykkyyttä suhteellisen piankin. Tie kohti tavoitetta osoittautui kuitenkin pidemmäksi ja vaikeammaksi kuin liki 70 vuotta sitten luultiin.[1] Matkan varrella maalia on siirretty – ihmisälyn kaltainen älykkyys on edelleen kaukana, mutta pitkälti tekoälytutkimuksen ansiosta meillä on muun muassa hakukoneita, joiden ansiosta tieto on saatavilla helpommin kuin koskaan, kasvojentunnistuksella aukeavia älypuhelimia, lähes itsenäisesti ajavia autoja ja nopeaa rokotekehitystä vastaamassa pandemioiden uhkiin.

1 AI 100 2021: Michael L. Littman, Ifeoma Ajunwa, Guy Berger, Craig Boutilier, Morgan Currie, Finale Doshi-Velez, Gillian Hadfield, Michael C. Horowitz, Charles Isbell, Hiroaki Kitano, Karen Levy, Terah Lyons, Melanie Mitchell, Julie Shah, Steven Sloman, Shannon Vallor & Toby Walsh. *"Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report."* Stanford University, Stanford, CA, September 2021. Haettu: 16.9.2021.

Tekoälyä on hyvin vaikea määritellä, eikä yhtä yleisesti hyväksyttyä määritelmää tiedeyhteisön piirissä olekaan pystytty muodostamaan. Perusongelma on se, että luonnollista älykkyyttäkään ei ole onnistuttu määrittelemään yleispätevästi. Tässä raportissa keskitytään pitkälti koneoppimiseen, joka onkin tärkeä, joskaan ei ainoa, tekoälytutkimuksen osa-alue. Koneoppiva tekoäly on **keinotekoinen järjestelmä, joka kykenee tekemään havaintoja ympäristöstään ja muuttamaan tai säätämään toimintaansa tekemiensä havaintojen perusteella** – järjestelmä on siis adaptiivinen eli jossain mielessä ”oppiva”, vaikka tällainen adaptiivisuus ei ihmisen oppimista paljon muistutakaan.

Purkamalla määritelmää hieman pääsemme syvemmälle sen sisältöön: Keinotekoisen järjestelmän ydin on nykyratkaisuissa tietokoneohjelma. Ohjelma toimii tietokoneessa, ja se voi saada tietoa ympäristöstä antureillaan, esimerkiksi kameralla. Se voi vaikuttaa ympäristöönsä, esimerkiksi suoraan ohjaamalla robottia tai välillisesti avustamalla ihmistä päätöksenteossa. Oppiva tekoälyohjelma muuttaa itse säätöarvojaan tai koodiaan niin, että sen tuottamat tulokset palvelevat paremmin sille annettua tavoitetta. Ohjelman toimintaympäristö voi olla puhtaasti digitaalinen tai se voi olla konkreettinen, kuten tehdas tai autotie.

Tekoäly voidaan karkeasti jakaa *heikkoon* (suppeaan) ja *vahvaan* (yleiseen) tekoälyyn. Heikko tekoäly on se, jota käytännössä on jo pystytty toteuttamaan, vahva tekoäly puolestaan on tavoite, johon voidaan pitemmällä aikavälillä pyrkiä. Vahva tekoäly on myös paljon esillä tieteiskirjallisuudessa ja -elokuvissa. Näitä ei ole syytä ylenkatsoa, koska niiden avulla päättäjät, kansalaiset ja tutkijat voivat pohtia haluttuja ja ei-toivottuja kehityssuuntia. Tällöin on kuitenkin hyvä pitää mielessä, että asiantuntijat ovat hyvin erimielisiä siitä, kuinka kauan kestää ennen kuin ensimmäisiä vahvoja, ”oikeasti älykkäitä” tekoälyjärjestelmiä pystytään rakentamaan, jos koskaan. On joka tapauksessa hyvin epätodennäköistä, että sellaista ihmisälykkyyttä muistuttavaa vahvaa tekoälyä, joka voisi muodostaa ihmisille jonkinlaisen uhan, voisi syntyä vahingossa. Kaikki nykyiset ja vielä pitkällä tulevaisuudessakin kehitettävät tekoälyjärjestelmät perustuvat heikkoon ja kapeaan tekoälyyn, mutta ne voivat olla ihmisille suunnattoman hyödyllisiä työkaluja. Toinen tapa jaotella tekoälyä on esitetty Tekoälyn kokonaiskuva ja kansallinen osaamiskartoitus -selvityksen loppuraportissa (2019)[2], jossa tekoäly jäsennetään kymmeneksi osaamisalueeksi data-analyysista robotiikkaan.

Nykyään käytössä olevat tekoälyratkaisut perustuvat hyvin suurelta osin koneoppimiseen, jossa siis tekoälyohjelmat kykenevät oppimaan datasta ja säätämään toimintaansa. Koneoppiminen poikkeaa perinteisestä ohjelmoinnista, jossa kaikki muutokset tekee ohjelmoija, joka miettii jokaisessa tapauksessa yksityiskohtaisesti, miten ohjelman on reagoitava eri tilanteissa. Koneoppimisessa ohjelmoijan ei tarvitse ratkaista jokaista yksityiskohtaa itse, vaan riittää, kun hän luo alustavat raamit ja mekanismin, jonka avulla ohjelma voi sitten datan perusteella tietyissä suppeissa rajoissa muokata itse itseään. Koneoppimisella pystytään kuitenkin luomaan vain hyvin yksinkertaisia ohjelmia, joiden tehtävä on kapea ja hyvin määritelty, ja lopputuloksena on aina suppean tekoälyn sovellus.

---

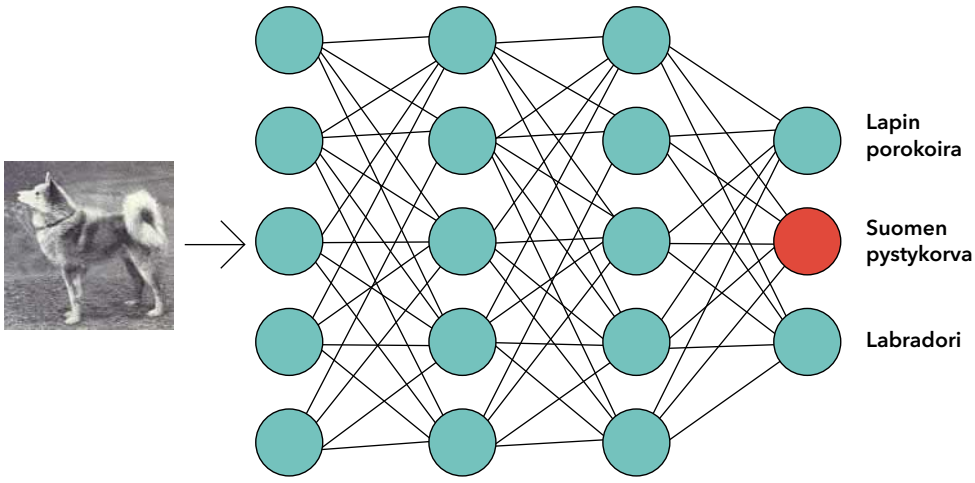
2 Tekoälyn kokonaiskuva ja kansallinen osaamiskartoitus – loppuraportti. Ailisto, Heikki; Neuvonen, Anssi; Nyman, Henrik; Halén, Marco; Seppälä, Timo (2019). Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 4/2019. Haettu 4.10.2021.

Koneoppiminen soveltuu erityisen hyvin esimerkiksi aikasarjojen ennustamiseen ja hahmontunnistukseen. Aikasarja voi olla esimerkiksi Suomen sähkönkulutus ensi viikon aikana. Hahmontunnistuksesta hyvä esimerkki on hienoisen poikkeaman havaitseminen potilaan röntgenkuvasta.

On huomattava, että tällaisen tehtävän ratkaiseva tekoälyohjelma ei yksinään muodosta hyödyllistä tietokonejärjestelmää, vaan se on aina osa laajempaa modulaarista järjestelmää, jossa osa moduuleista saattaa kerätä tekoälyohjelman tarvitsemaa dataa, osa suorittaa toimenpiteitä tai laatia toimenpidesuosituksia tekoälyohjelman tulosten perusteella ja osa huolehtia koneen ja ihmisen käyttöliittymästä. Tekoälyjärjestelmäkin on siis loppujen lopuksi aina ihmisten rakentama tietokonejärjestelmä, josta suurimman osan muodostavat perinteisellä tavalla ohjelmoidut moduulit mutta jossa yksi tai useampi kapean ja hankalasti ohjelmoitavan tehtävän hoitava moduuli rakennetaan tekoälymenetelmiä käyttäen. Koska näitä moduuleja ei rakenneta kokonaan käsin, keskeisiä kysymyksiä ovat tekoälymoduulien ymmärrettävyys ja niiden toiminnan oikeellisuuden ja luotettavuuden validointi.

1.1.1 Koneoppiminen neuroverkoilla

Koneoppivien menetelmien kenttää hallitsevat nyt keinotekoiset neuroverkot, erityisesti niin sanotut syvät neuroverkot. Keinotekoinen neuroverkko on etäisellä tavalla ihmisen aivosoluja muistuttava laskentamalli, joka on toteutettu tietokoneohjelmana. Neuroverkko koostuu suuresta määrästä alkeellisia laskentaosasia, ”neuroneja”, jotka liittyvät toisiin neuroneihin (kuva 1). Neuronit on järjestetty kerroksiksi, jotka muodostavat eräänlaisen putken tai liukuhihnan, jossa edellisen kerroksen ulostulo on seuraavan syöte. Käytössä neuroverkon alkupäähän syötetään data, esimerkiksi koiraa esittävän kuvan pikselit, ja syöte etenee kerros kerrokselta liukuhihnan läpi tuottaen toisessa päässä toivottavasti halutun tuloksen, esimerkiksi tunnistuksen kuvan aiheesta: ”tämä on suomenpystykorva”. Syväoppivaksi tai syväksi neuroverkkoa sanotaan, kun siinä olevien neuronien ja kerrosten määrä on hyvin suuri, ja verkko sisältää esimerkiksi satojatuhansia, miljoonia tai jopa miljardeja neuroneja.



This photo by Unknown Author is licensed under CC BY-SA

KUVA 1. Neuroverkko kuvantunnistuksessa.

Neuroverkon varsinainen ydin on opetusalgoritmi, joka opetusvaiheen aikana muokkaa kussakin alkeellisessa laskentaosasessa (neuronissa) olevia kertoimia vähitellen niin, että neuroverkko antaa oikeita tuloksia. Opetusvaiheessa neuroverkolle syötetään opetusdataa, esimerkiksi kuvia erirotuisista koirista, ja samalla kerrotaan, mikä on oikea tulos. Opetuksen aikana neuroverkon tunnistustarkkuus kasvaa, ja se kykenee tunnistamaan oikein myös syötteitä, joita se ei ole koskaan aiemmin saanut, esimerkissämme uuden kuvan koirasta. On syytä huomata, että koneoppiminen saattaa tuottaa tilastollisesti hyviä tuloksia (esimerkiksi 95 prosenttia vastauksista oikein), mutta sen antama tulos ei ole koskaan absoluuttinen totuus.

1.1.2 Koneoppimisen haasteet

Neuroverkoilla on saavutettu näyttäviä tuloksia viimeisten kymmenen vuoden aikana esimerkiksi kuvien tunnistamisessa ja luonnollisen kielen tuottamisessa. Edistykseen voidaan nähdä neljä syytä:

- tietokoneiden kasvanut laskentateho mahdollistaa suuremmat ja siksi paremmin toimivat neuroverkot,
- neuroverkkojen opettamiseen on saatavissa runsaasti hyvälaatuista dataa,
- neuroverkkojen ohjelmointityökalut ovat kehittyneet erittäin paljon, jolloin ne ovat muidenkin kuin alan tutkijoiden hyödynnettävissä,
- koneoppimisalgoritmit ovat kehittyneet tutkimuksen ansiosta paremmiksi.

Nykyiset tekoälyjärjestelmät ovat datapohjaisia. Tällä tarkoitetaan sitä, että ne oppivat datasta, siis tavallaan ohjelmoivat itseään datan pohjalta. Siksi laadukkaan ja runsaan datan saatavuus on niille välttämätöntä.

On vielä syytä korostaa, että nykyisillä tekoälyohjelmilla ei ole minkäänlaista tietoisuutta eikä todellista ymmärrystä käsittelemästään informaatiosta. Informaatio ja data ovat tekoälyohjelmalle vain numeroita, joita se käsittelee algoritminsa ohjaamana. Tekoälyohjelmat ovat tästä huolimatta suuri edistysaskel tavallisiin tietokoneohjelmiin verrattuna, ja niiden tuottamat hyödyt ja niiden aiheuttamat muutokset ovat hyvin suuria monilla talouden ja yhteiskunnan aloilla.

Muutoksen taustalla on yksinkertaisesti se havainto, että koneoppimisalgoritmeilla voidaan luoda melko helposti ohjelmistoja sellaisiin tehtäviin, joiden ratkaiseminen perinteisellä ohjelmoinnilla olisi ollut työlästä tai mahdotonta. Kyseiset ohjelmat eivät ole missään nimessä älykkäitä vahvan tekoälyn merkityksessä, mutta kapea tekoäly riittää mainiosti lukemattoman moniin hyödyllisiin sovelluksiin, esimerkiksi älypuhelimen käyttäjän tunnistamiseen kuvan perusteella tai erottamaan läpivalaisukuvissa samanlaisia poikkeamia kuin lääkärit.

Miten sitten tästä eteenpäin, yhä suurempiin neuroverkkoihin ja yhä vaativampiin tehtäviin? Tekniseltä kannalta yhä monimutkaisemmat neuroverkot pystyvät ratkaisemaan kapeita, hyvin määriteltyjä tehtäviä kaiken aikaa paremmin, mutta tämä ei välttämättä tarkoita sitä, että se on oikea tie kohti yleiskäyttöisempää tai vahvempaa tekoälyä: yksi tekoälyohjelma osaa laskea go-pelin seuraavan siirron niin, että sen mukaan pelaava amatööri voittaa parhaan go-ammattilaisen, mutta sama tekoälyohjelma ei osaa itse

siirtää nappuloita tai havaita peliasetelmaa – puhumattakaan siitä, että se tietäisi pelaavansa peliä (tai edes sitä, että on olemassa ihmisiä). Samaa tekoälyohjelmaa ei voi soveltaa go-pelin lisäksi mihinkään muuhun peliin, saati sitten vaikkapa auton ohjaamiseen tai läpivalaisukuvan tunnistamiseen. Neuroverkkojen parametrien lukumäärän kasvattamien satoihin tai tuhansiin miljardeihin – mikä tällä hetkellä on jo todellisuutta – ei suoraan johda yhtään yleiskäyttöisempään tai vahvempaan tekoälyyn vaan on omiaan entisestään pahentamaan seuraavia neuroverkkojen kriittisiä ongelmia:

- Koska neuroverkkojen toiminta perustuu miljoonien tai miljardien parametrien tuottamiin ”mikrovaikutuksiin”, kokonaisuuden hahmottaminen on vaikeaa. Neuroverkkojen toimintaa ei ole yleensä mahdollista yleisellä tasolla selittää tai summata kovinkaan ymmärrettävästi (vaikkakin niiden rakentamismekanismi on sinänsä hyvin ymmärrettävä, ja itse asiassa matemaattiselta kannalta varsin yksinkertainen).
- Parametrien lukumäärän lisääminen mahdollistaa entistä suurempien data-aineistojen hyödyntämisen, mutta tällöin etulyöntiaseman, jopa tietynlaisen monopoliaseman, saavat sellaiset teknojätit, jotka ovat osanneet kerätä tällaisia valtavia datamassoja, usein yksityisyyden suojan kustannuksella.
- Neuroverkkojen säätämiseen eli ”opettamiseen” tarvittavat laskentaresurssit kasvavat parametrien ja datan lukumäärän kasvaessa ja suosivat osapuolia, jotka osaavat parhaiten muuttaa keräämänsä datan rahaksi esimerkiksi personoidun mainonnan kautta. Samaan aikaan koneoppimiseen käytetyt energiamäärät alkavat olla globaalisti jo hälyttävällä tasolla, vaikka tekoälyvallankumous ei ole laajamittaisesti vielä edes alkanut.

Seuraavan sukupolven tekoälyjärjestelmistä on rakennettava tehokkaampia, niin että ne pääsevät hyvään tulokseen pienemmällä määrällä dataa, laskentaa ja energiaa. Niiden pitää toimia laaja-alaisemmin, mahdollisesti heterogeenisiä datalähteitä yhdistäen, ja kehittyä ihmisten toimintaa tehostaviksi hyödyllisiksi, monipuolisiksi ja luotettaviksi kokonaisuuksiksi. Tekoälyjärjestelmien ei siis tule pelkästään automatisoida erillisiä, aikaisemmin ihmisten tekemiä yksinkertaisia tehtäviä.

1.2 Tekoäly muutosvoimana

Maailman talousfoorumi WEF näkee tekoälyn neljännen teollisen vallankumouksen ajurina[3], ja OECD puolestaan korostaa tekoälyä laajasti yhteiskuntaan vaikuttavana murroksellisena voimana[4]. Muutos ei rajoitu talouteen, vaan se koskee myös laajasti julkista hallintoa, terveydenhuoltoa, maanpuolustusta ja kansallista turvallisuutta sekä kansalaisyhteiskuntaa. Murrokseen liittyy luonnollisesti sekä suuria mahdollisuuksia että uhkia.

3 Maailman talousfoorumin tekoälysivusto.
4 OECD:n tekoälysivusto.

Ei olekaan yllättävää, että kansallisvaltiot, EU, kansainväliset organisaatiot ja suuryritykset ovat laatineet tekoälystrategioita. Suomi oli ensimmäisten joukossa julkistaessaan kansallisen ohjelman *Suomen tekoälyaika*[5].

Tekoälyn merkitys näkyy investoinneissa: suuryritykset sijoittivat tekoälyteknologiaan vuonna 2020 noin 60 miljardia euroa[6], ja tekoäly-yrityksiin sijoitettiin vuonna 2019 noin 36 miljardia euroa[7]. Näiden lisäksi tulevat kansalliset investoinnit tutkimukseen ja tekoälyn sotilas-, valvonta- ja turvallisuussovellutuksiin. Tekoälyn hyödyntäminen on myös muodostunut suurvaltojen kilpakentäksi, jolla Yhdysvallat ja Kiina ovat vahvassa asemassa.

Suomen lähtökohdat tekoälyteknologioiden hyödyntämiseen ovat hyvät. Perustan muodostavat kansainvälisestikin merkittävä tieteellinen tutkimus ja koulutus, joiden juuret ulottuvat 1960- ja 1970-luvuille, sekä Suomen yleinen vahvuus digitaalisuudessa. Lisäksi kansalaisten asenteet ja kyvykkyys uuden teknologian käyttöönottoon ovat eurooppalaisittain hyvällä tasolla. Viime vuosina maahamme on syntynyt vahvojen, kansainvälisillä markkinoilla toimivien yritysten rinnalle kasvuyritysten ja start-upien joukko. Suomen potentiaali on noteerattu konsulttiyrityksissä[8], jotka ennakoivat Suomen olevan tekoälystä eniten hyötyvien kansantalouksien joukossa esimerkiksi Yhdysvaltojen ja Ruotsin ohella.

Tekoälyn hyödyntämiseksi ja uhkien torjumiseksi EU on luomassa tekoälyn käytön harmonista sääntelyä (*Artificial Intelligence Act*), ja samantyyppistä lainsäädäntöä luodaan myös muissa maissa. OECD[9] on julkistanut tekoälyn käyttöä ohjaavat periaatteet, ja vastaavia kannanottoja on muodostettu esimerkiksi YK:ssa.

On kuitenkin syytä huomata, ettei tekoälyteknologia ole vielä lunastanut sille asetettuja suuria talouskasvuun liittyviä odotuksia. Esimerkiksi Accenture[10] ennusti vuonna 2017 tekoälyteknologioiden hyödyntämisen kaksinkertaistavan BKT:n kasvun perusuraan verrattuna Yhdysvalloissa ja Suomessa vuosina 2017–2030, niin että vuotuinen kasvu olisi neljä prosenttia kahden sijasta. Vastaavasti konsulttiyritys McKinsey arvioi Tekoälyaika-raportissa (2017) siteeratussa Suomea koskevassa selvityksessään tekoälyn tuottaman lisäkasvun olevan samat kaksi prosenttia perusuraan verrattuna. Ennustettu tekoälyn vauhdittama tuotannon ja tuottavuuden kasvu ei ole ainakaan toistaiseksi toteutunut. Toisaalta myöskään tekoälyyn liitetyt uhkakuvat massatyöttömyydestä eivät ole realisoituneet.

---

5 Suomen tekoälyaika (2017): *Suomen tekoälyaika – Suomi tekoälyn soveltamisen kärkimaaksi: Tavoite ja toimenpidesuositukset.* Tekoälyohjelman ohjausryhmä. Työ- ja elinkeinoministeriön julkaisuja 41/2017.

6 Statista (2021): Global total corporate artificial intelligence (AI) investment from 2015 to 2020.

7 **Brookings Instituten analyysi tekoälyinvestointien kehityksestä**.

8 Purdy M. ja Daugherty P. 2017. *Why artificial intelligence is the future of growth.* Accenture; Suomen tekoälyaika (2017).

9 **OECD:n tekoälyperiaatteet**.

10 Purdy M. ja Daugherty P. (2017): *"Why artificial intelligence is the future of growth".* Accenture.

Vaikka tekoäly ei ole vielä siivittänyt maailman ja Suomen taloutta tuottavuuslähtöiseen suureen kasvuun, tämä ei tarkoita, etteikö tekoälyteknologioiden soveltamisella ole jo nyt suuria vaikutuksia yrityksissä ja yhteiskunnissa. Voi myös olla niin, että uuden teknologian laajamittaisempi leviäminen ja merkittävimmät vaikutukset tulevat pitemmällä viipeellä kuin oletettiin, koska nykyisten perusmenetelmien suoraviivainen soveltaminen ei edellä mainittujen teknisten ongelmien vuoksi riitä tekoälyn lopulliseen läpimurtoon kaikilla yhteiskunnan osa-alueilla. Tämä tarkoittaa sitä, että tekoälyn jatkokehitykseen ja erilaisten teknologisten tai yhteiskunnallisten pullonkaulojen poistamiseen vahvimmin investoivat tahot tulevat hyötymään tästä murroksesta jatkossa eniten. Lähtökohdat ovat Suomella ja Euroopalla oikein hyvät, mutta kisa on vasta alussa.

1.3 Viestit päättäjille ja kansalaisille

Tekoälyteknologioiden tutkimuksella, kehittämisellä ja hyödyntämisellä on suuri merkitys Suomen tulevalle kilpailukyvylle sekä kansalaisten hyvinvoinnille ja turvallisuudelle. On syytä varmistaa, että Suomi pystyy käyttämään hyvät lähtökohtansa tekoälyn hyödyntämisessä.

Koska tekoälyssä kyse ei ole yksittäisestä teknologiasta vaan monitahoisesta menetelmäperheestä, se on alana nopeasti kehittyvä ja varsin laaja. Päättäjien tulisikin varata aikaa aihealueen kehityksen ja siitä käytävän keskustelun seuraamiseen ainakin yleisellä tasolla.

Erityisesti on tärkeää vaikuttaa aktiivisesti EU:n tekoälyä koskevaan sääntelyyn, niin että se luo Eurooppaan selkeän, ymmärrettävän ja innovaatiopositiivisen toimintaympäristön. Vahingossakaan ei pidä antaa vaikutelmaa ylimitoitetuista, raskaista ja monimutkaisista prosesseista, minkä myötä yritykset tai julkiset toimijat jättäisivät tekoälyratkaisut hyödyntämättä. Luonnollisesti tekoälyteknologioiden käytön on oltava sopusoinnussa yhteiskunnan perusarvojen ja kansalaisten oikeuksien kanssa, minkä pitääkin olla kaiken eurooppalaisen tekoälysääntelyn vahvana taustamotiivina.
Konkreettisiksi tavoitteiksi kannattaa asettaa:

- perus- ja soveltavan tutkimuksen resurssien vahvistaminen,
- yritysten ja tutkimusmaailman tiiviin yhteistyön tukeminen,
- tekoälyteknologioiden määrätietoinen ja strategisesti johdettu hyödyntäminen julkisella sektorilla terveydenhuollosta maanpuolustukseen,
- tekoälyn huippuosaajien riittävä kouluttaminen ja rekrytoinnit ulkomailta sekä osaajien sitouttaminen.
- täydennys- ja jatkokoulutuksen sekä työssäoppimisen lisääminen kehittämään tekoälytaitoja, joiden ansiosta kansalaisten työmahdollisuudet laajentuvat ja hallinnan tunne työssä parantuu.

Kansalaisten on syytä seurata keskustelua tekoälyn käytöstä, vaikutuksista ja eettisistä kysymyksistä ja osallistua siihen. Kansalaisille tarkoitetut tekoälyyn liittyvät verkkokurssit, esimerkiksi Helsingin yliopiston ja Reaktorin *Tekoälyn perusteet*, tarjoavat erinomaisen mahdollisuuden tutustua asiaan pintaa syvemmältä. Ne auttavat hälventämään joitakin tieteiskirjallisuudesta tuttuja pelkoja ja tuovat toisaalta realismia tekoälyä kohtaan asetettuihin suuriin odotuksiin.

2 Tekoäly tehostaa ja analysoi

Sasu Tarkoma
professori
Helsingin yliopisto

Tekoäly on keskeinen digitaalisen infrastruktuurin rakennusosa ja merkittävä tekijä Suomen kilpailukyvyn kannalta. Käynnissä on murros, joka liittyy uuteen data- ja tekoälyvetoiseen toimintamalliin ja alustatalouden kasvuun. Uuden mallin mukaiset organisaatiot vievät tekoälyn soveltamisen useiden, elleivät kaikkien, prosessien läpi ja saavat siitä merkittävää kilpailukykyetua ja arvon lisäämistä. Datassa nähdään valtava lisäarvo, mutta datan saatavuus ja laatu sekä tietosuoja ja datavetoisen automaation eettinen pohja herättävät paljon kysymyksiä, joita tutkitaan aktiivisesti Suomessa ja kansainvälisesti. Tekoälyjärjestelmien sääntely yhtäältä varmistaa toiminnan eettistä pohjaa, toisaalta monimutkainen ja teknologiaspesifi sääntely voi rajoittaa innovaatioiden syntymistä ja käyttöä. Laajemmin ohjelmistojen eettiseen toimintaan ja mallikkuuteen liittyvä sääntely ja digitaalisen infrastruktuurin osittainen desentralisaatio kehittyvät toisiaan tukien, jolloin lähestytään datan ketterää ja eettistä hyödyntämistä suuressa mittakaavassa. Tulevaisuuden hajautetun, reunalaskentaan pohjautuvan tekoälyn ansiosta tietoa voidaan analysoida eettisesti ja turvallisesti siellä, missä analyysia tarvitaan.

2.1 Mistä on kyse?

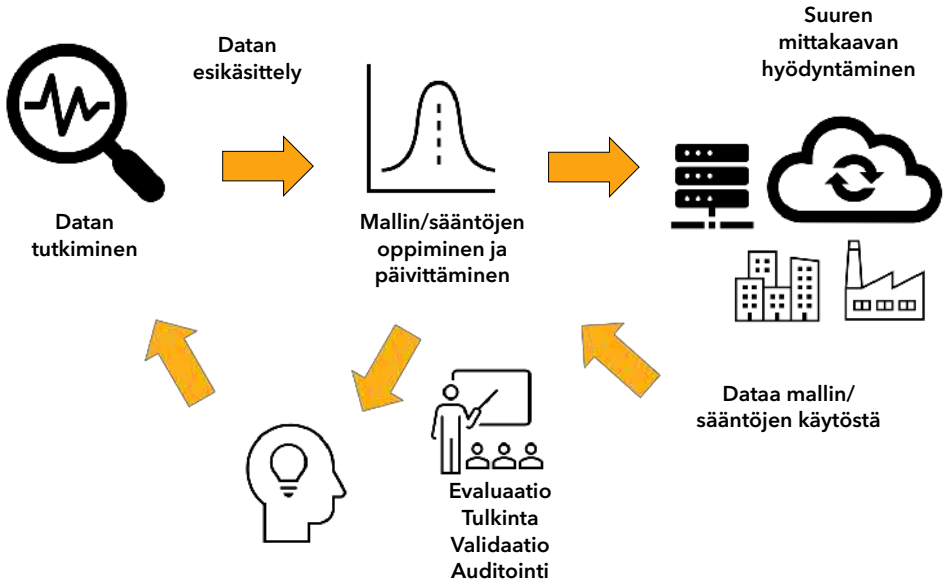
Tekoälyjärjestelmät ovat ohjelmisto- ja laitteistojärjestelmiä, jotka annetun monitahoisen tavoitteen saavuttamiseksi havainnoivat ympäristöä, hankkivat dataa ja tekevät dataan ja tietoon perustuvia päätelmiä ja päätöksiä[1]. Tekoälyjärjestelmät keräävät ja käsittelevät sekä jäsenneltyä että jäsentämätöntä dataa ja tekevät päätelmiä tästä datasta johdetusta tietämyksestä. Tekoälyjärjestelmät rakentuvat useiden eri menetelmien varaan. Niistä keskeisiä ovat koneoppimismenetelmät sekä symbolinen tekoäly. Tekoäly voi perustua symbolisiin sääntöihin, tai se voi oppia numeerisen mallin datasta. Tekoälyjärjestelmä voi tutkia ja analysoida aiempien päätöstensä vaikutuksia toimintaympäristöön ja mukauttaa käyttäytymistään sen perusteella.

Käytettyjen menetelmien perusteella tekoälyjärjestelmät voidaan jakaa kahteen ryhmään: sääntöpohjaisiin järjestelmiin ja oppiviin järjestelmiin. Sääntöpohjaisten järjestelmien toimintaa voidaan selittää ja testata säännöstön avulla. Niiden päätöksenteon läpinäkyvyys ja selitettävyys ovat saavutettavissa olevia ominaisuuksia. Oppivissa järjestelmissä päättelyn ja lopputuloksen selitettävyys voi olla haasteellisempaa saavuttaa. Oppiva järjestelmä on riippuvainen sen koulutukseen käytetystä aineistosta ja syötemateriaalista.

1 Euroopan komissio. *A Definition of AI: Main Capabilities and Scientific Disciplines.* Tekoälyä käsittelevä korkean tason asiantuntijaryhmä (AI HLEG), 2018.

Erityisesti neuroverkkojen ja niihin perustuvan syväoppimisen alueilla toiminnan ja päätöksenteon läpinäkyvyyttä ja selitettävyyttä pidetään ongelmallisina.

Viime vuosina käytettävissä oleva laskentakyky sekä algoritmit ovat kehittyneet merkittävästi ja tekoälyn tarvitsemaa dataa on enemmän kuin ennen. Tulevaisuuden kvanttitietokoneet voivat tuoda merkittävää lisätehoa laskennallisten ongelmien ratkaisuun ja simulaatioiden toteuttamiseen. Varsinaista tekoälymenetelmiin liittyvää paradigman muutosta ei kuitenkaan ole tapahtunut, eikä sellaista ole näköpiirissä. Paradigma on kuitenkin muuttunut datavetoisessa toiminnassa, mikä on johtanut globaalin alustatalouden syntymiseen.



KUVA 1. Tekoälyjärjestelmien kehittämisen yleinen prosessi.

Kuva 1 havainnollistaa tekoälyjärjestelmien kehittämisen yleistä prosessia. Prosessissa keskitytään erityisesti tekoälymalleihin ja säännöstöihin. Tekoälyn kehittäminen edellyttää ilmiön ja toimintaympäristön mallintamista, jossa usein tarvitaan ihmisen tulkintaa ja ohjausta. Koneoppiminen keskittyy mallien rakentamiseen, päivittämiseen ja käyttöön. Tuotettuja malleja hyödynnetään päätöksenteossa ja ilmiöiden ennustamisessa. Tekoäly voi myös oppia jatkuvasti ja päivittää mallia, kun uutta dataa on saatu. Tämä toimintatapa on hyödyllinen teollisten järjestelmien ohjauksessa ja esimerkiksi autonomisten järjestelmien toteuttamisessa. Vahvistusoppimisessa robotti tai autonominen järjestelmä tutkii ympäristöään ja oppii toimimaan toimintaympäristössä positiivisen ja negatiivisen palautteen perusteella.

Tekoälyjärjestelmissä voi tapahtua useita erilaisia vääristymiä, jotka liittyvät tekoälyn kehittämiseen ja käyttöön. Vääristymä voi olla tahallinen tai tahaton, ja aiheuttaa myönteisen tai kielteisen seurauksen. Vääristymä voi aiheuttaa syrjivän lopputuloksen. Sääntöperusteisten tekoälyjärjestelmien vääristymä voi syntyä säännöstön määrittelyssä

tai käytössä. Oppivien järjestelmien vääristymä voi tapahtua tietojen keräämisessä ja tekoälyn koulutuksessa. Jatkuvasti oppivissa järjestelmissä vääristymä voi syntyä esimerkiksi verkossa tapahtuvan oppimisen ja vuorovaikutuksen kautta. Vääristymien tunnistamiseen tarvitaan useita eri menetelmiä ja menettelyitä, kuten asiantuntijoiden käyttöä, algoritmeja ja tarkistusmekanismeja. Sekä sääntöpohjaiset että oppivat ratkaisut vaativat testausta ja auditointia, jotta oikeellisuus voidaan varmistaa. Kuva 1 havainnollistaa tätä datan ja kehitettävän tekoälyjärjestelmän arviointia ja tarkistusta. Yleistä automaattista menetelmää vääristymien tunnistamiseen tai tekoälyjärjestelmän auditointiin ei ole.

Tekoäly on keskeinen digitaalisen infrastruktuurin rakennusosa ja merkittävä tekijä kansallisen ja EU:n kilpailukyvyn kannalta. EU:n komission huhtikuussa 2021 julkaisema tiedonanto (COM(2021) 205 final) kiteyttää eurooppalaisen lähestymistavan tekoälyyn. Komissio edistää ihmiskeskeistä tekoälyä, joka rakentuu ihmisoikeuksien, vahvan eettisen pohjan ja tietosuojan varaan. EU tähtää tekoälyteknologian hyödyntämiseen kokonaisvaltaisesti. Tekoälyn tulee olla lainmukainen, eettinen sekä teknisesti ja sosiaalisesti luotettava[2].

2.2 Tekoäly muutosvoimana

Elämme kiihtyvän digitalisaation maailmassa, jossa yhteiskunnan toimintaympäristö on muuttunut ja muuttuu nopeasti globaalin kilpailun ja kiihtyvän automatisaation yhteisvaikutuksesta. Maailma on suurten haasteiden edessä. Näitä ovat esimerkiksi pandemia ja siitä toipuminen, ilmastonmuutos ja informaatiovaikuttaminen. Ilmastonmuutos ja puhdas luonto vaativat nopeita ja tehokkaita päätöksiä niin globaalilla tasolla kuin jokapäiväisessä elämässäkin. Kehittyvät tekoälymenetelmät ja ilmiöistä kerättävä data mahdollistavat entistä paremmat tiedot päätöksentekoa ja suurten haasteiden käsittelyä varten. Tämä onnistuu kuitenkin vain, jos data ja siitä jalostettava tieto ovat luotettavia ja tietoa voidaan ottaa vastaan ja käsitellä päätöksenteon eri prosesseissa.

Alustataloudesta on muodostunut nykymaailman keskeinen kilpailutekijä, joka rakentuu pilvilaskennan ja modernin tietoliikenneinfrastruktuurin varaan. Taulukko 1 tiivistää tekoälyn keskeisiä sovellusalueita: teollisuus, julkinen sektori, kuluttajasektori, terveys ja hyvinvointi ja tiede. Teollisuudessa on käynnissä digitalisaation murros, jossa monipuoliset anturit, robotit, tietoverkot, pilvilaskenta ja tekoäly yhdistyvät turvallisiksi ja skaalautuviksi ratkaisuiksi. Tekoäly- ja optimointimenetelmillä on keskeinen asema näissä ratkaisuissa. Internet ja sen kuluttajille suunnatut palvelut ovat johtaneet nykytilanteeseen, jossa massadata ja sen käsittely pilvialustoilla ovat keskittyneet suurille alustatoimijoille. Niistä valtaosa on EU-alueen ulkopuolella. Tekoälymenetelmät auttavat toteuttamaan skaalautuvat tietoverkot sekä pilvipalvelut, ja ne räätälöivät palvelut ja sisällöt. Lähitulevaisuudessa tekoäly alkaa yhä enemmän tuottamaan sisältöä – esimerkiksi Netflix on jo tuottanut elokuvan tekoälyn tukemana.

---

2 Euroopan komissio. **Luotettavaa tekoälyä koskevat eettiset ohjeet**. Tekoälyä käsittelevä korkean tason asiantuntijaryhmä (AI HLEG). 2019.

Alue	Datalähteet	Mahdollisuudet
Valmistava teollisuus	**Data materiaalivirroista, tuotteista, roboteista, tehtaista ym.**	**Älykäs automaatio, diagnosointi, prediktiivinen huolto, digitaalinen kaksonen**
Julkinen sektori	**Avoin data, prosessit ja palvelut, kaupunkidata**	**Uudet palvelut, räätälöidyt palvelut, parantuneet prosessit**
Kuluttajasektori	**Data tuotteista ja älytiloista/ kaupungista, älykkäät kaupat, sosiaalinen media**	**Markkinointi ja mainonta, sentimenttianalyysi, personoidut palvelut, parantunut käyttäjäkokemus**
Terveys ja hyvinvointi	**Data potilaista ja ympäristöstä**	**Personoitu lääketiede, ennaltaehkäisevä hoito, parantuneet prosessit**
Tiede	**Hyvin erilaisia datalähteitä**	**Suurten datamäärien kautta lähestyttävät ongelmat, simulaatio**

Taulukko 1 Esimerkkejä tekoälyn sovellusalueista

Tekoälymenetelmät suodattavat ja tuottavat digitaalista sisältöä yhä automaattisemmin. Samalla ne luovat tilaisuuden myös haitalliselle vaikuttamiselle. Globaalin alustatalouden informaatiovirtojen avulla vaikutetaan ihmisten käyttäytymiseen ja yritysten toimintaan mikro- ja makrotasoilla. Kun vuorovaikutteista sisältöä tuottava teknologia kehittyy, on ensiarvoisen tärkeää varmistua sisällön alkuperästä ja oikeellisuudesta.

Tekoäly on yhteiskunnan eri alueiden läpileikkaava muutosvoima, jonka ansiosta voidaan tehostaa palveluita ja prosesseja. Paradigman muutoksen myötä voidaan pitemmällä aikajänteellä parantaa myös vuorovaikutteisia digitaalisia palveluita ja ratkaisuja. Tekoälyjärjestelmien kehittäminen ja hyödyntäminen pitävät sisällään myös merkittäviä riskejä, jotka liittyvät kansakuntien kilpailukykyyn, turvallisuuteen ja toiminnan eettisen pohjan ja tietosuojan varmistamiseen. EU:n tekoälyasetus pyrkii varmistamaan oikeudenmukaisen ja eettisen toimintaympäristön, mutta teknologian riskipohjainen sääntely voi myös heikentää innovaatioiden kehittämistä ja käyttöönottoa.

2.3 Viestit päättäjille ja kansalaisille

Käynnissä on teknologinen murros, joka liittyy uuteen data- ja tekoälyvetoiseen toimintamalliin ja alustatalouden kasvuun. Tekoäly mahdollistaa dataan ja tietoon pohjautuvan automatisaation, joka voi johtaa merkittävään kestävän tuottavuuden kasvuun.

Kilpailukyvyn kannalta keskeisiä teemoja ovat

- osaamispohjan kehittäminen ja osaajien saatavuus
- pitkäjänteinen tutkimus- ja kehittämistoiminta lähtien perustutkimuksesta
- datan saatavuus ja laatu
- suurien datamäärien käsittely pilvialustoilla ja pilvialustojen hajauttaminen
- etiikkaan, tietosuojaan ja tietoturvaan liittyvät kysymykset.

EU:n tekoälyasetus asettaa reunaehdot tekoälyjärjestelmien eettiselle käytölle, ja se tulee globaalisti vaikuttamaan tekoälyjärjestelmien kehittämiseen. Kilpailukyvyn kannalta olennaista on, että sääntely ei keskity muuttuviin menetelmiin ja teknologiaan vaan niiden soveltamiseen ja käyttöön. Kokonaisturvallisuuden kannalta olennaista on, että tekoälyteknologian sääntely mahdollistaa turvallisuusuhkiin varautumisen. Tarvitaan yleistä laaja-alaista osaamista tekoälyn yhteiskunnallisista vaikutuksista. Lisäksi tarvitaan syvällistä osaamista digitaalisen infrastruktuurin ja digitaalisten palveluiden kehittämiseksi, ylläpitämiseksi ja suojaamiseksi.

Organisaatioiden johtamis- ja toimintamalleja pitäisi arvioida siitä näkökulmasta, millaisia tekoälyn mahdollisuuksia niissä voidaan hyödyntää. Perinteisen organisaation toimintatapoja on haastavaa päivittää kerralla tekoälyä hyödyntäviksi. Hyväksi havaittu lähestymistapa on soveltaa tekoälyä liiketoimintamalliin pohjautuvaan toiminnan tehostamiseen. Sekä organisaation johtamisen että tuotteiden ja palveluiden toteuttamisen kannalta on olennaista, että organisaation sisäiset ja ulkoiset rajapinnat tukevat datan jakamista ja hyödyntämistä. Datan ymmärtäminen, yhdistäminen ja tuotteistaminen lisäävät merkittävästi kilpailukykyä ja avaavat uusia markkinoita[3].

Teknologian näkökulmasta tekoälyjärjestelmät ovat tyypillisesti osa suurempaa ohjelmistoihin ja laitteistoon perustuvaa järjestelmää, jota päivitetään ja kehitetään usein modernien ohjelmistoprosessien mukaisesti. Tarvitaan uusia menetelmiä ja prosesseja jatkuvasti kehitettävien ja tekoälyä hyödyntävien ratkaisuiden ylläpitoon, suojaamiseen ja toiminnan varmistamiseen. Tekoälyjärjestelmien kehittäminen vaatii paitsi dataa myös riittävän laskentaympäristön ja tekoäly- ja alustaosaamista.

Varteenotettavana vaihtoehtona nykyiselle, pitkälti keskitetylle pilvipalvelu- ja alustatalousmallille on kehittymässä uusi reunalaskennan paradigma, jossa datan paikallinen käsittely korostuu. Reunalaskennassa laskeminen on tuotu lähelle dataa tuottavia ja käyttäviä laitteita. Tällöin on mahdollista käsitellä entistä enemmän paikallista dataa, datan käsittelyn viiveet pienenevät ja vain olennainen tieto välitetään eteenpäin, esimerkiksi perinteiselle pilvialustalle. Kun reunalaskenta yhdistetään hajautettuihin tekoälymenetelmiin, voidaan entistä paremmin hyödyntää paikallista dataa. Näin varmistetaansamalla, että myös tietosuojaan ja tietoturvaan liittyvät vaatimukset tulee otettua huomioon.

---

3 Kirjallisuutta: Marco Iansiti ja Karim R. Lakhani. *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Boston: Harvard Business Review Press, 2020.

Suomelle tärkeät teemat, tekoäly ja langaton tietoliikenne, yhdistyvät 5G:tä seuraavassa 6G:ssä. Tulevaisuuden hajautettu, reunalaskentaan pohjautuva tekoäly mahdollistaa tiedon analyysin ja käsittelyn eettisesti ja turvallisesti siellä, missä analyysia tarvitaan. Suomella on merkittävä asema ja mahdollisuus näyttää EU:lle tietä tekoälyn eettisessä, turvallisessa ja tehokkaassa käytössä.

3 Tekoäly havainnoi – koneaistit

Joni-Kristian Kämäräinen
signaalinkäsittelyn professori
Tietotekniikan yksikkö
Tampereen yliopisto

3.1 Tekoäly havainnoi – tulin, näin, voitinko?

Tekoälyä käytetään kaikissa teknologioissa, joita tarvitaan sellaisten robottien rakentamiseen, jotka pystyvät tekemään fyysisiä töitä. Robotti havainnoi koneaisteilla, kuten kuulo-, näkö- ja tuntoaisteilla, liikkuu raajoilla, telaketjuilla, pyörillä tai lentäen ja vaikuttaa ympäristöönsä koskettamalla ja tarttumalla. Tällaisen robotin pitäisi kyetä tekemään ihmisen sille antamia tehtäviä, mikä vapauttaisi ihmiset tärkeämpiin tehtäviin. Tällainen tärkeämpi tehtävä voi olla vaikkapa toisen ihmisen, asukkaan, kohtaaminen seniori-talossa.

Kun robottien ja muiden koneiden havainnointia on toteutettu, keskeisiä ovat olleet tekoäly ja erityisesti koneoppimismenetelmät. Erittäin merkittävässä roolissa ovat olleet aisteista kenties tärkein, näkö, ja sen toteuttaminen laskennallisesti eli tietokonenäkö (engl. *computer vision*). Nyt on sopiva hetki välitilinpäätökselle, jossa tarkastellaan, mistä tultiin, missä ollaan ja mihin pitäisi mennä seuraavaksi.

3.2 Koneaistit – tietokonenäkö tekoälyn veturina

Tekoälyä on tutkittu aktiivisesti jo 1970-luvulta lähtien, jolloin järjestettiin ensimmäiset tekoälyä koskevat tieteelliset konferenssit ja alettiin julkaista ensimmäisiä sitä koskevia tieteellisiä lehtiä. Koska ihmisen älykkyyttä vastaavan laskennallisen älykkyyden kehittäminen on houkutteleva tutkimuskohde, ala houkutteli ja houkuttelee edelleen tutkijoita eri tieteenaloilta.

Tietokonenäköä ja erityisesti kohteiden tunnistamista pidettiin yhtenä avaintekijänä tekoälyn kehittämiseen. Kohteiden tunnistaminen (engl. *object detection*) oli merkittävin soveltavan tekoälytutkimuksen haaste ja se lopulta johtikin murrokseen, kun Professori Li Fei-Fein tutkimusryhmän keräämä ImageNet-kuva-aineisto otettiin käyttöön. Tietokonenäön tutkimus onkin houkutellut valtavasti tutkijoita 90-luvun lopusta alkaen. Tietokonenäöstä tuli tekoälyalan suurin tutkimusala tutkijoiden ja artikkeleiden määrällä mitattuna, ja sen pääkonferenssi, CVPR, hipoo viittausmääriltään sellaisia tieteen kärkilehtiä kuin *Science* ja *Nature*.

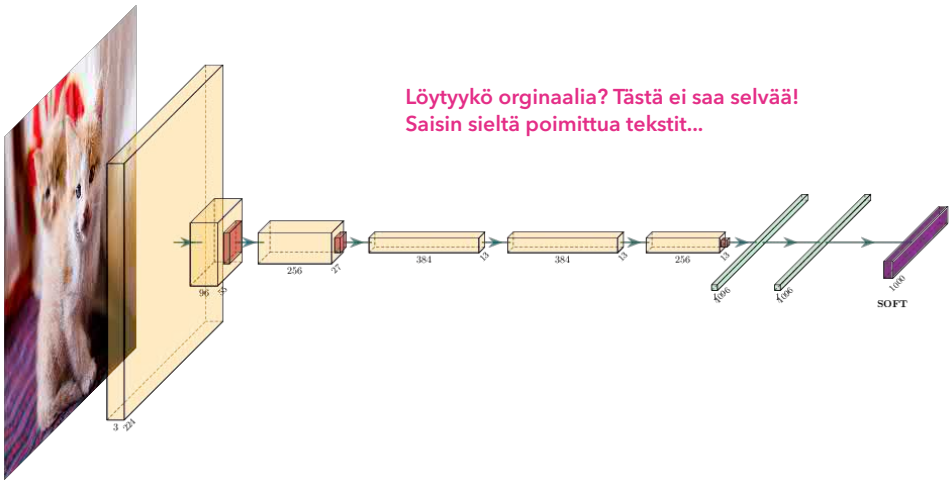
Tietokonenäössä tavoitteena on rakentaa kameralla varustetulle tietokoneelle ihmisen näköä vastaava näkökyky. Vastaavasti konekuulossa (engl. *machine hearing*) tavoitteena on rakentaa mikrofonilla varustetulle tietokoneelle ihmisen kuuloa vastaava kuuloaisti. Tietokonenäön ja muiden koneaistien tutkimus kehittyi tasaisesti vuoteen 2012 asti.

Tuona vuonna brittiläis-kanadalaisen Toronton yliopiston professorin Geoffrey Hintonin (s. 1947) tutkimusryhmä julkaisi läpimurtoartikkelin[1]. Artikkelissa esiteltiin uusi lähestymistapa, joka voitti tietokonenäön kehittämisen seuratuimman kilpailun, *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)*. Vuotta 2012 voidaankin pitää modernin tekoälyn syntymävuotena.

Hintonin ryhmän menetelmä perustui laskennalliseen rakenteeseen nimeltä *konvoluutioneuroverkko* (engl. *convolutional neural network, CNN*). Konvoluutioneuroverkko eli lyhyemmin *konvoluutioverkko* oli uusi laskentaparadigma, jossa uusien menetelmien kehitys tehdään korkeammalla ja siksi myös yksinkertaisemmalla tasolla. Kehitystyössä täytyy yhdistellä erilaisia laskentakerroksia ongelmaan sopivaksi arkkitehtuuriksi, ja verkko oppii loput opetusaineistosta eli datasta. Pascal VOC -kilpailussa opetusaineiston muodostivat kuvat ja niiden luokat. Opittu tieto tallentuu kerrosten painoihin (engl. *weights*), joita Hintonin ryhmän arkkitehtuurissa on kaikkiaan 60 miljoonaa.

Hintonin ryhmän esittelemä verkko, *AlexNet*, on esitetty kuvassa 1. Verkon painoja opetetaan siten, että miljoonia opetuskuvia syötetään toistuvasti verkon läpi. Verkon toinen pää tuottaa kullekin kuvalle tuhat todennäköisyyttä, joista jokainen edustaa yhtä luokkaa (esim. ”Auto”, ”Kissa” tai ”Koira”). Ihannetilanteessa vain oikean luokan todennäköisyys on 100 prosenttia ja kaikkien muiden 0. Aluksi näin ei kuitenkaan ole, vaan ulostulon virhe ohjataan uudelleen verkon läpi ja jokaisen kerroksen painoja säädetään virheen pienentämiseksi.

Hintonin ryhmän AlexNet-verkko mullisti tekoälytutkimuksen. Nykyään kaikkien koneaistien aloilla parhaat tulokset saadaan konvoluutioverkoilla.



KUVA 1. Konvoluutioverkko koostuu kerroksista. Kun kuva kulkee kerrosten läpi, kerrokset tuottavat kunkin luokan todennäköisyyden. Kuva vastaa Hintonin ryhmän vuonna 2012 esittelemää arkkitehtuuria, AlexNetiä.

---

1 A. Krizhevsky, I. Sutskever, G. E. Hinton (2012): *ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems* (NeurIPS).

3.2.1 Yksi menetelmä löytää heidät, se yksi heitä hallitsee

Ennen konvoluutioverkkojen kehittämistä tekoälytutkimus oli sirpaloitunut pienempiin tutkimusyhteisöihin. Tietokonenäön, luonnollisen kielen käsittelyn ja äänen analysoinnin tutkijat kävivät eri konferensseissa ja julkaisivat eri lehdissä, ja heidän käyttämänsä menetelmät olivat erilaisia.

Vuoden 2012 jälkeen kaikki tekoälytutkimuksen alat ovat siirtyneet käyttämään konvoluutioverkkoihin perustuvia arkkitehtuureja. Tämän seurauksena alojen väliset erot ovat kaventuneet. Nyt tietyn alan tutkija voi ottaa nopeasti käyttöönsä parhaan mahdollisen menetelmän joltakin toiselta alalta ja jopa parantaa sitä. Vuosi vuodelta konvoluutioverkkojen käyttö laajenee yhä useammille tutkimus- ja sovellusalueille, kuten biologiaan ja lääketieteeseen[2].

Konvoluutioverkoista onkin tullut kuin Sormusten sormus Taru Sormusten herrasta -teoksessa: ”Yksi sormus löytää heidät, se yksi heitä hallitsee.”

3.3 Tekoälyn nykytila – kun Suomi putos puusta

Vielä 1990-luvulla Suomi oli tekoälyn kärkimaa ja täällä tehtiin juuri neuroverkkotutkimusta. Kiitos tästä kuuluu akateemikko Teuvo Kohoselle (s. 1934) ja hänen kouluttamilleen tutkijasukupolville. 2000-luvulla neuroverkkotutkimuksesta luovuttiin laajalti ja keskityttiin muihin koneoppimisen menetelmiin.

Neuroverkkotutkimuksen pariin palattiin Suomessa ja erityisesti Aalto-yliopistossa 2010-luvun lopulla. Sitä on myös ansiokkaasti yhdistetty 2000-luvulla tehtyyn tutkimukseen, kuten bayesilaiseen päättelyyn. Neuroverkkotutkimuksen uudessa nousussa ovat auttaneet merkittävät panostukset Aalto-yliopiston tekoälytutkimukseen. Aalto-yliopisto onkin taas merkittävä tutkimuslaboratorio. Aalto-yliopiston tutkijat ovat tehneet myös uusia avauksia, kuten konvoluutioverkkojen käytön tietokonegrafiikassa.

Koneaistien alalla Suomessa tehdään merkittävää tutkimusta ainakin tietokonenäössä ja konekuulossa. Tietokonenäön tutkimus on aktiivisinta Aalto-yliopistossa sekä Tampereen ja Oulun yliopistoissa. Konekuulon tutkimus puolestaan on merkittävää Aalto-yliopistossa ja Tampereen yliopistossa. Muiden aistien tutkimus on Suomessa vähäistä. Suomeen olisikin hyvä saada laajempi teknologinen osaaminen koneaisteissa, kuten tuntoaistissa (engl. *tactile sensing*), joka on tärkeä robotiikassa.

Tekoälytutkimuksen tärkeimmällä osa-alueella eli konvoluutioverkoissa perustutkimus on vahvaa Aalto-yliopistossa, jossa on saavutettu kriittinen tutkijoiden määrä. Muiden menetelmien ja koneaistien tutkimuksessa riittävän tutkijoiden määrän on saavuttanut useampi yliopisto: Helsingin yliopisto, Tampereen yliopisto ja Oulun yliopisto. Nämä yliopistot pystyvät vastaamaan Suomen koulutus-, tutkimus- ja tuotekehitystarpeisiin, ja

---

2 P. Gainza et al. (2020): *Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning.* Nature Methods 17.

panostuksia onkin hyvä suunnata jo valmiiksi vahvoihin yliopistoihin. Muissa yliopistoissa on hyvä painottaa paikallisia tarpeita, koska vahvaa teollisuutta on niiden läheisyydessä laajemmin.

Tekoälyyn, koneoppimiseen ja koneaisteihin suuntautuvaa tutkimusta ja opetusta pitää vahvistaa, koska kaikki ennusteet viittaavat siihen, että näillä aloilla on tulevaisuudessa merkittävä rooli. Paras toimintatapa on, että vahvistetaan kunkin yliopiston vahvuusaluetta, koska resurssit eivät riitä tekoälyn perustutkimuksen nostamiseen riittävän tutkijamäärän yli jokaisessa suomalaisessa yliopistossa.

3.3.1 Tutkimus siirtynyt yrityksiin – kaiken takana on data, joka ihanasti opettaa

Edellä on puhuttu vain menetelmistä ja erityisesti konvoluutioverkoista, mutta tekoälyalalla on tapahtunut myös toinen, paljon merkittävämpi muutos, jonka konvoluutioverkot ovat mahdollistaneet. Konvoluutioverkkomenetelmien toimivuus nimittäin riippuu niiden opetuksessa käytetyn datan määrästä ja laadusta. Tätä rajoitetta on yritetty kiertää muun muassa luomalla simuloitua dataa, mutta tähänastiset kokemukset siitä ovat olleet heikkoja. Näyttääkin siltä, että mikään ei voita laadukasta aitoa dataa.

Koska data on menetelmiä tärkeämpää ja datan keräämistä harvoin pidetään tieteellisenä työnä, monen tutkimusongelman kannalta paras osaaminen on luonnollisesti siirtynyt yrityksiin. Yrityksillä on resursseja kerätä ja merkitä (engl. *annotate* tai *label*) valtavia määriä dataa, esimerkiksi kuvia, ääntä ja lääketieteellisiä näytteitä.

Vaarana on, että yritykset vaikuttavat siihen, mitä tutkitaan, koska ne voivat jo pelkällä tutkijoidensa määrällä ohjata sitä, minkä ongelmien ratkaisemiseen keskitytään ja mitkä kuuluvat marginaaliin. Tärkeimmissä tekoälyalan konferensseissa merkittävä osa tiedeartikkeleista tulee suurimmista yrityksistä. Jopa yksittäinen yritys pystyy kilpailemaan suurienkin maiden kanssa tiedeartikkelien määrässä ja näin ohjaamaan tieteen kehitystä.

Suurin osa tekoälytutkimuksesta ja erityisesti tietokonenäön ja konekuulemisen tutkimuksesta on siirtynyt jättiyrityksiin. Näitä ovat muun muassa Google, Facebook, Apple, Amazon, Microsoft, Baidu, Alibaba ja Huawei. Lisäksi laskenta-alustojen yksipuolisuus eli näytönohjainteknologiaan perustuvien GPU-korttien käyttö on nostanut Nvidian erittäin merkittävään asemaan. Nvidian aseman ja sirupulan vuoksi monilla suomalaisilla tutkijoilla on ollut vaikeuksia saada laskentaan soveltuvia GPU-kortteja.

Esimerkkinä datan voimasta voidaan mainita sähköautojen valmistaja Tesla, jonka itseajava ohjelmisto on johtava maailmassa. Teslan tekoälyjohtaja Andrej Karpathyn mukaan hänen johtamansa ryhmä käyttää 75 prosenttia ajastaan datan keräämiseen ja merkitsemiseen. Todellinen osuus saattaa kuitenkin olla jopa suurempi. Teslan menetelmän kehitys perustuu siis datan hankkimiseen, ja vastaavia yrityksiä tulee varmasti lisää.Suomi ja Eurooppa ovat suhtautuneet dataan sinisilmäisesti, minkä vuoksi ne ovat menettäneet kilpailuasemansa datan keräämisessä ja hyödyntämisessä. Tähän on yritetty vaikuttaa EU-tason päätöksillä, mutta ne eivät ole jarruttaneet datan kulkeutumista suuryrityksille. Jotkin päätökset ovat kärjistäen olleet idealistisen herttaisia mutta käytännössä hyödyttömiä. Eurooppa tarvitsee vähemmän herttaisia ja aiempaa vaikuttavampia ja itsekkäämpiä päätöksiä.

Suomen ja Euroopan pitää tehdä toimia eurooppalaisen datan suojelemiseksi niin, että eurooppalaisilla toimijoilla, yrityksillä ja yhteiskunnalla on siihen ensisijainen oikeus kansainvälisiin toimijoihin nähden. Lisäksi on syytä ylläpitää osaamista laskenta-alustojen tutkimuksessa ja tuotekehityksessä, esimerkiksi *System on Chip-* eli *SoC-teknologioissa*, koska on nähtävissä, että tietoturvalliset, tiettyyn tarkoitukseen suunnitellut ja tietoturvahyökkäyksiä sietävät sirut ovat merkittävässä roolissa tulevaisuuden autonomisissa laitteissa. Esimerkiksi Tesla on päätynyt tekemään omia siruja konvoluutioneuroverkkolaskentaan.

3.4 Tehdään Suomesta taas suuri – oppia Elokapinalta ja Gretzkyltä

Tekoälyssä kilpailu maiden ja yritysten välillä lisääntyi voimakkaasti 2010-luvun jälkipuoliskolla, ja 2020-luvulla kilpailu on ollut erittäin kovaa. Lähes kaikilla mailla on oma tekoälystrategiansa, ja kaikki maat tekevät mahdollisuuksiensa mukaan massiivisia investointeja tekoälyn tutkimukseen ja tuotekehitykseen. Tekoälyalan yrityksille, kuten kaikille digitaaliajan yrityksille, on ominaista niin sanottu voittaja saa kaiken -periaate. Se tarkoittaa, että lopulta jää jäljelle vain yksi tai muutama yritys, jotka hallitsevat markkinoita. Näin syntyy monopoliasema, joka on haitaksi muille yrityksille ja maille.

3.4.1 Eurooppalainen vastarintaliike, datakapina

Seuraavassa kuvataan lyhyesti suosituksia Euroopan tason toimiksi.

Suositus 1: Datamonopolien tuhoaminen suuryrityksiä pilkkomalla
Suomi ja Eurooppa ovat datateollisuuden kehitysmaita, ja kuluttajadataa hyödyntävillä markkinoilla onkin vaikeaa saada jalansijaa. Tämä johtuu pääasiassa siitä, että tietyt kansainväliset yhtiöt ovat kasvaneet liian suuriksi[3]. Euroopan pitäisi vaatia näiden yhtiöiden pilkkomista niin, että dataa tuottavat osat, esimerkiksi Googlen hakukone, eriytetään omaksi liiketoiminnakseen, joka tarjoaa palveluitaan kaikille yrityksille samoilla kaupallisilla ehdoilla. Näin estettäisiin monopoliaseman väärinkäyttö.

Suositus 2: Eurooppalainen kuolontähti (datamonopoli)
Jos Eurooppa ei kykene vaatimaan suuryritysten pilkkomista, toinen vaihtoehto on ottaa käyttöön Kiinan ja Venäjän malli ja rakentaa omat jätit Eurooppaan. Kiina ja Venäjä lähtivät jo varhaisessa vaiheessa kilpailemaan yhdysvaltalaisia yrityksiä vastaan, ja tuloksena oli omat hakupalvelut Baidu ja Yandex, jotka ovat kasvaneet suuryrityksiksi ja ekosysteemeiksi. Eurooppa voisi rakentaa yhteisellä rahoituksella omat kilpailevat yrityksensä ja vielä vaikuttaa markkinoihin. Euroopan etuna ovat luotettavuus ja avoimuus. ”Euroopan google” olisi myös Eurooppaa yhdistävä toimi, joka houkuttelisi Eurooppaan osaajia ja investointeja.

Toinen mahdollisuus on kopioida Eurooppaan yhteiseen hyvään pyrkivän *OpenAI*-yrityksen konsepti. Yhtiö on ideologiallaan houkutellut monia hyviä tutkijoita, ja se on myös tuottanut useita läpimurtoja. Odotettavissa on, että sillä on lopulta myös kaupallista merkitystä.

---

3 S. Galloway (2017): The Four: The Hidden DNA of Amazon, Apple, Facebook, and Google.

3.4.2 Pienen Suomen pitää tehdä sitä, missä se on hyvä, eikä sitä, mikä on kivaa

Suositus 3: Tukea datan keräämiselle
Suomalaisia ja eurooppalaisia yrityksiä pitäisi tukea siinä, että ne tunnistavat, mitä mahdollisesti kilpailuetua tuovaa dataa niillä on tai mitä ne voisivat helposti kerätä. Suomen Akatemian (SA) ja Business Finlandin (BF) rahoitusvälineet eivät tue datan keräämistä, koska sitä ei pidetä perustutkimuksena (SA) tai sen tuomaa arvoa ei osata mitata (BF). Siksi tulisi kehittää tutkimus- ja tuotekehitysvälineitä, joissa ensisijaisena ehtona olisi maailman mittakaavassa laadullisesti ja määrällisesti erityisen data-aineiston kerääminen.

Suositus 4: Raskas teollisuus – rikas, rämä, mutta oma
Suomen kokoisen maan ei kannata kilpailla kaikilla tekoälyn osa-alueilla. Suomen vahvuuksia ovat raskas teollisuus, kuten raskaat työkoneet, haastava ilmasto ja autonomiset laitteet ääriolosuhteissa sekä korkeatasoiset sosiaalipalvelut, kuten terveydenhuolto sekä palvelut lapsille ja senioreille. Tutkimus ja tuotekehitys näillä aloilla vaatii tietokoneiden lisäksi laitteita, laboratorioita ja osaamista, jotka pudottavat heikoimmat kilpailijamaat kisasta. Suomen ei kannata kilpailla pelkässä datan prosessoinnissa, koska siihen tarvitaan infrastruktuuriksi vain tietokoneita ja se on siksi mahdollista ja houkuttelevaa köyhimmillekin maille. Siinä määrä voikin voittaa kyvykkyyden. On paljon hyödyllisempää kehittää kalliita ja raskaita laitteita, joiden avulla pudotetaan monia kilpailijoita pois markkinoilta. Tällaisia ovat työkoneet, laivat, lentoalukset, tehdashallien ja satamien nosturit ja hissit sekä raskas robotiikka. Kun keskitytään näitä aloja palvelevaan tekoälyn ja koneaistien tutkimukseen, suomalainen insinöörityö ja kokemus pystyvät kilpailemaan kansainvälisesti.

3.4.3 Mihin kiekko menee seuraavaksi? Luonnollinen kieli ja robotiikka

Tekoälyn murros alkoi tietokonenäöstä, mutta nyt ala on saturoitunut valtavasta yritystutkimuksen määrästä. Saattaa olla, että seuraava ”tekoälyn talvi” (engl. *AI winter*) alkaa juuri tietokonenäöstä. Uudet teknologiat on kuitenkin nopeasti omaksuttu myös muilla aloilla. Onkin ehkä syytä noudattaa jääkiekkoilija Wayne Gretzkyn isän neuvoa: ei kannata mennä sinne, missä kiekko on nyt, vaan sinne, mihin se menee seuraavaksi.

Kaksi nopeasti nousevaa mahdollista läpimurtoalaa ovat

- robotiikka ja
- luonnollinen kieli.

Robotit ovat odottaneet tulemistaan koteihin ja pienyrityksiin jo vuosikymmeniä, mutta nyt on odotettavissa, että tekoälyn avustamana murros viimein tapahtuu. Suomessa merkittävää robotiikkatutkimusta tehdään käytännössä enää Aalto-yliopistossa ja Tampereen yliopistossa. Näissäkin ollaan jäljessä Ruotsia ja Tanskaa. Näiden yliopistojen robotiikan tutkimusta pitäisi vahvistaa. Pitäisi myös harkita jopa robotiikan yksiköiden ja koulutusohjelmien perustamista.

Luonnollisen kielen prosessointi (engl. *natural language processing, NLP*) oli tärkeässä roolissa tekoälyn alkuvaiheessa, kunnes 2000-luvulla se ajautui marginaaliin ja sen tutkimus ja koulutus hävisivät monista yliopistoista. Konvoluutioneuroverkot ovat kuitenkin nostaneet kielen prosessoinnin uuteen kukoistukseen, ja nyt osaajista on valtava pula. Tämän alan tutkimuksella on pitkät perinteet ainakin Helsingin ja Turun yliopistoissa, mutta myös muualla. Yliopistoihin ja yrityksiin saataisiin osaajia investoimalla välittömästi alan koulutukseen ja tutkimukseen.

4 Tekoäly liikkuu ja tekee

Juha Röning
professori
Oulun yliopisto

Robotit vaikuttavat kaikkialla. Teollisuudessa ne lisäävät kilpailukykyä, yhteiskunnallisesti ne tuovat palveluja terveyden ja hyvinvoinnin alalle sekä kuljetukseen, turvallisuuteen ja elintarviketuotantoon, ja taloudellisesti niillä on vaikutusta kaikilla aloilla. Robottitekniikka vaikuttaa voimakkaasti yhteiskunnan hyvinvointiin ja moniin eri ympäristöihin, ei pelkästään valmistavaan teollisuuteen. Samalla täytyy pitää mielessä peruslähtökohta: Robotit ovat fyysisiä. Ne toimivat vuorovaikutuksessa ja käsittelevät, liikuttavat ja muotoilevat osia, tavaroita sekä ihmisiä.

Suomen täytyy ylläpitää markkinaosuutensa nopeasti laajentuvassa maailmanmarkkinatilanteessa. Tämä edellyttää voimakkaita toimenpiteitä ja jatkuvaa hereilläoloa. Robotiikkaa voidaan nyt käyttää monilla reaalimaailman aloilla, valmistuksen lisäksi auttamaan ihmisiä työssä ja terveydenhuollossa, siirtämään tavaroita ja ihmisiä ja tarkastamaan ja ylläpitämään rakenteita.

International Federation of Robotics kerää vuosittain tilastot ja kokoaa katsauksen robottien käytöstä maailmanlaajuisesti. Perinteisesti raportissa on käsitelty teollisuusrobotteja, mutta nyttemmin näiden rinnalle ovat nousseet erilaiset palvelurobotit ja niiden sovellukset.

Teollisuusrobotit ovat monikäyttöisiä ja yleisesti ohjelmoitavia sekä automaattisesti ohjattuja robotteja. Niissä on vähintään kolme niveltä, ja niitä käytetään teollisissa automaatiosovelluksissa. Perinteisesti ne on jaoteltu viiteen eri kinemaattiseen tyyppiin, jotka on varustettu käyttötapauskohtaisilla päätelaitteilla.

Teollisuusrobottien ja nimenomaan automaation kysyntä on kasvanut vuodesta 2010 lähtien. Vuosina 2014–2019 kasvua on tullut yli 10 prosenttia vuodessa. Vuonna 2018 asennettiin jo 400 000. yksikkö. Maailman suurin teollisuusrobottimarkkina-alue on selkeästi Aasia. Vuodesta 2012 lähtien suurin markkina on ollut Kiina, jonka osuus on 38 prosenttia kaikista asennuksista. Euroopassa isoin markkina-alue on Saksa. Teollisuusrobottien tärkein asiakas on autoteollisuus, jonne asennetaan lähes kolmannes kaikista teollisuusroboteista.

Palvelurobotit suorittavat erilaisia tehtäviä. Ne on yleensä suunniteltu käyttökohteen mukaan, ja niissä on useimmiten vähemmän kuin kolme niveltä. Palvelurobotti ei yleensä ole täysin itsenäinen, vaan toimii kauko-ohjauksella.

Palvelurobotit voidaan jakaa ammattipalvelurobotteihin ja kotitalouspalvelurobotteihin. Kotitalouksien palvelurobottien lukumäärä on monikertainen verrattuna teollisuusrobottien lukumäärään. Vuonna 2020 niitä oli myyty jo liki 27 miljoonaa yksikköä. Suurelta

osin tämä selittyy edullisilla imurityyppisillä kotiroboteilla. Ammattipalvelurobotteja (erityisesti logistiikka ja sairaalateknologia) asennettiin samaan aikaan 250 000 yksikköä. Ammattikäyttöön tarkoitetut laitteet ovat huomattavasti arvokkaampia investointeja, joten markkinan kokonaisvolyymissa ne olivat liki neljä kertaa kotitalouksien investointeja suurempia. Palvelurobotteja kehitetään eniten niin yritysten kuin tuotteidenkin määrällä mitattuna Euroopassa. Koronapandemia näyttää aktivoineen palvelurobottien kehitystä ja markkinoita toisin kuin teollisuusrobottien kehitystä ja markkinoita. Viimeisen vuoden aikana markkinoille on esimerkiksi tullut yli 30 erilaista desinfiointirobottia.

Raja teollisuus- ja palvelurobottien välillä on häilyvä. Sovelluksesta riippuen sama yksikkö voi olla palvelurobotti tai teollisuusrobotti. Myös robottien käytettävyys on muuttumassa. Uusia sovelluksia kehitetään jatkuvasti ja niitä pystytään entistä paremmin myös toteuttamaan käytännössä. Robotti ei ole enää häkillä eristetty teollisuusautomaatti, vaan se mahdollistaa täysin uudenlaista liiketoimintaa.

Tällä hetkellä puhutaan paljon yhteistyörobotista eli kobotista (engl. *cobot*). Tekoäly- ja koneoppimistekniikoiden myötä robotit pystyvät aistimaan ja havainnoimaan ympäristöään ja vastaamaan interaktiivisesti sen muutoksiin. Robotti tukee yhä enemmän ihmisen toimintaa, ja siitä on tullut työkaveri, joka jakaa saman työskentelyalueen.

Nostan tässä raportissa esille kaksi uutta robotiikkasovellusten toimialaa normaalin teollisen tuotannon lisäksi: terveysteknologian palvelurobotit ja -sovellukset sekä erilaiset teolliset ja muut tarkastustoimet. Näiden kehittäminen ei onnistu ilman, että ympäristöstä on reaaliajassa saatavilla tietoa, jonka pohjalta voidaan ymmärtää erilaiset tilanteet ja sovittaa robotin toiminta niihin.

4.1 Tekoäly muutosvoimana: Tekoälyn, datan ja robotiikan mahdollisuudet

4.1.1 Robotiikan tutkimus Euroopassa

Robotiikka nousi uuteen kukoistukseen EU:n edellisen puiteohjelmakauden eli Horisontti 2020:n aikana. Sen myötä tutkimukseen kohdennettiin 700 miljoonan euron rahoitus. Komissio siirtyi myös PPP-malliin, jossa julkinen (*public*, komissio) ja yksityinen (*private*, tutkimus) keskustelevat keskenään ja ohjaavat yhdessä (*partnership*, euRobotics) tutkimusta oikeaan suuntaan.

Tällä periaatteella sovittiin myös maailman suurin siviilirobotiikan T&K&I-ohjelma SPARC, jonka kokonaispanostus koostui 0,7 miljardin euron rahoituksesta EU:lta vuosina 2014–2020 ja 2,1 miljardin euron yksityisrahoituksesta teollisuuden puolelta.

SPARCin toiminta pohjautui kahteen tiekartta-asiakirjaan. Strategic research agenda (SRA) oli korkean tason strateginen suunnitelma, joka tarjosi yleiskatsauksen laajalle lukijakunnalle ja määritteli toiminnan terminologian, kun taas multiannual roadmap (MAR) tarkensi teknisiä yksityiskohtia. MARia oli tarkoitus päivittää joka vuosi, jolloin tutkimuksessa voitaisiin seurata nopean kehityksen tuomia uusia trendejä ja mahdollisuuksia. Mielestäni tämä saatiin toimimaan hyvin, vaikka alussa esiintyikin joitakin lähinnä aikataulutuksiin liittyviä haasteita.

Puolivälin tarkastelussa alkoi jo olla selvää, että seuraavaan puiteohjelmaan saadaan ja otetaan Horisontti 2020:stä hyviä eväitä, mutta robotiikan kehitykseen tarvittiin selvästi entistä laajempaa yhteistyötä. Tältä pohjalta syntyi tarve yhdistää entistä vankemmin data ja sen ymmärtäminen toiminnanohjaukseksi analyysin, tekoälyn ja oppivien menetelmien avulla.

Tätä työtä on nyt realisoitu perustamalla uusi yhteistyöorganisaatio Adra[1], joka yhdistää toiminnassaan tekoälyn (CLAIRE, ELLIS, EurAI), datan (BDVA) ja robotiikan (euRobotics). Adran tavoitteena on turvata Euroopan suvereniteetti tekoäly-, data- ja robotiikkateknologioiden osaamisessa, vahvistaa Euroopan johtavaa asemaa tekoäly-, data- ja robotiikkateknologioissa, joilla on suuri sosioekonominen vaikutus, ja tukea Euroopan vahvaa maailmanlaajuista kilpailuasemaa tällä alalla.

Adran menestyksen periaate on avoimuus ja osallisuus. Laaja ja tasapainoinen vuoropuhelu on välttämätöntä, jotta voidaan yhdistää kaikki eurooppalaisen innovaatioekosysteemin osa-alueet mukaan lukien start-upit, pk-yritykset, suuret teollisuusyritykset sekä tutkimuslaitokset ja hyödyntää täysimääräisesti tekoälyn, datan ja robotiikan eurooppalaiset vahvuudet.

4.1.2 Tekoäly – hyvä renki, mutta huono isäntä

Robotit ovat alkaneet siirtyä tehdassalin liukuhihnalta entistä monipuolisempiin tehtäviin. Enää ne eivät välttämättä ole sidottuja yhteen paikkaan, vaan ne kykenevät liikkumaan, tai ne voidaan joustavasti sijoittaa eri tehtäviin.

Robotti ei myöskään enää vain käytä konetta, vaan sen sallitaan työskennellä myös ihmisen lähellä. Kehitys kulkee osaltaan kohti yhteistyörobotteja, mikä tarkoittaa sitä, että robotti ja ihminen jakavat saman työskentelyalueen.

Robotiikan tutkimusalueella ollaan siirtymässä uuteen kauteen. Nykyinen, kahdeksas puiteohjelma Horisontti 2020 lähestyy loppuaan, ja Horisontti Eurooppa on aukeamassa. Horisontti 2020 oli robotiikan tutkimuksen ja kehityksen kannalta käänteentekevä, sillä tarjosihan se yrityksille ja tutkimuslaitoksille 700 miljoonaa euroa kilpailtavaa tutkimusrahaa. Vanhemmat tutkijat muistavat sellaisetkin puiteohjelmakaudet, jolloin *robotti*-sanan maininta hakuasiakirjoissa melkein aiheutti hakemuksen hylkäämisen. Komission kanta oli, että robotit on kehitetty valmiiksi, sillä nehän toimivat jo tuotantolinjoilla hyvin ja moitteetta.

Horisontti Euroopan valmistelu aloitettiin käytännössä silloin, kun nykyinen puiteohjelma ehti puoleenväliin. Silloin tehtiin merkittäviä linjauksia tulevaisuudentutkimuksen teemoista. Yksi isoista linjoista on ollut nostaa tekoäly tutkimuksen keskiöön. Tekoäly ajatellaan sateenvarjoksi, jonka alle mahtuu myös robotiikkatutkimus. Mielestäni asia ei ole näin yksioikoinen.

1 Adran lehdistötiedote 23.6.2021.

Tekoäly on perinteisesti keskittynyt täydelliseen automaatioon, jossa tietokone ratkaisee koko ongelman ilman ihmistä. Nyt kuitenkin näyttää siltä, että sen sijaan, että korvataan ihmiset, ongelmaa on parempi lähestyä suunnittelemalla järjestelmä, joka mahdollistaa ihmisen ja tekoälytyökalun yhteistoiminnan. Tehokkaassa yhteistyössä käytetään esimerkiksi luonnollisen kielen ymmärtämistä ja elekielen, aktiivisuuden ja aikomusten tunnistamista, jotta pysytään luomaan yhteisiä mentaalisia malleja ja suunnittelemaan interaktiivisuutta. Näitä tarvitaan, jos halutaan, että toimintaa ohjaa ihminen. Robotiikka voi ja sen tulee hyödyntää tekoälyä, mutta robotiikka on oma oppiaineensa, joka vaatii syvää osaamista. Fyysinen siirtyminen tai siirtäminen ei tapahdu ilman konetta ja mekaniikkaa.

Kestävyys, luotettavuus, turvallisuus ja ihmisoikeuksien noudattaminen muodostavat ne pääperiaatteet, joiden pohjalta roboteille[2] voidaan kehittää tekoäly- tai muita ohjelmistoja.

4.1.3 Tekoälyn mahdollisuudet robotiikassa

Kun noudatetaan edellä mainittuja periaatteita, tekoälyllä on robotiikassa laajat mahdollisuudet. Otan seuraavaksi kaksi esimerkkiä robotiikan mahdollisuuksista palvelusektorilla ja teollisessa tarkastamisessa.

Nykyään kuuluu entistä huolestuneempia äänenpainoja esimerkiksi hoitohenkilöstön jaksamisesta ja erityisesti sairaaloiden ja hoitokotien työvoimapulasta[3]. Kirjallisuudesta löytyy jo laaja valikoima terveydenhoidon robotiikkasovelluksia[4]. Esimerkiksi puhtaanapito on saanut suuren huomion koronapandemian ansiosta. Liikkuvat robotit voidaan varustaa UV-valolampuilla pintojen puhdistamiseksi. Samalla on otettava huomioon huoli työpaikkojen menettämisestä tai epäselvistä vastuukysymyksistä[5]. Terveydenhuollon palvelukseen pystytään valmistamaan entistä kyvykkäämpiä robotteja tekoälyn ja aiempaa monipuolisemman datan ansiosta.

Yksi kohde, jota voidaan kehittää robottien datan ja tekoälyn yhteistyöllä, ovat sosiaaliset robotit. Näitä yhteistyörobotteja voidaan tuoda terveydenhuollon eri tehtäviin antamalla niille tunnetietoisia ominaisuuksia, joilla mahdollistetaan ihmisystävällinen ja ihmistä avustava vuorovaikutus.

Eräs tärkeä robotiikan käyttökohde on infrastruktuurin huolto ja tarkastus. Robotiikan käyttöön ajavat infrastruktuurin omistajan velvollisuudet sekä turvallisuus-, ympäristö- ja taloudelliset vaikutukset. Suurin osa infrastruktuureista on tarkastettava säännöllisin väliajoin huoltotarpeiden tai turvallisuusvaatimusten vuoksi. Tarkastettavat rakenteet voivat sijaita vaarallisissa tai syrjäisissä paikoissa. Tarkastettavia kohteita on esimerkiksi

---

2 JRD conference and workshop report *What future for European robotics?* A science for policy perspective JRC virtual conference 27-29 January 2021.

3 Drennan, V. M. & Ross, F. (2019). *Global nurse shortages: The facts, the impact and action for change*. British medical bulletin, 130(1): 25–37.

4 Holland, J., Kingston, L., McCarthy, C., Armstrong, E., O'Dwyer, P., Merz, F., & McConnell, M. (2021). *Service robots in the healthcare sector.* Robotics, 10(1).

5 Cresswell, K., Cunningham-Burley, S., Sheikh, A., et al. (2018). *Health care robotics: qualitative exploration of key challenges and future directions.* Journal of medical Internet research, 20(7).

kuljetusalalla (rautatiet, kadut, sillat jne.), uusiutuvan energian, öljyn ja kaasun saralla, sähkönjakelussa, prosessiteollisuudessa (kemikaalit, lannoitteet jne.) sekä käyttö- ja jätevesijärjestelmissä (putket, kanavointi).

Erilaiset lentävät robotit (dronet, UAV:t) ovat vakiintuneet viimeisen vuosikymmenen aikana kiinnostavaksi tutkimusaiheeksi monilla eri alueilla, kuten mallinnuksen, ohjauksen, havainnoinnin ja jopa lentävällä laitteella tehtävän manipulaation saralla. Ne ovat myös tehokas teknologinen ratkaisu erilaisiin tehtäviin, kuten huoltoon ja tarkastukseen, etsintään ja pelastukseen, kuljetukseen ja tavaroiden toimitukseen sekä valvontaan ja 3D-kartoitukseen

Lentävät robotit ovat osoittautuneet hyödyllisiksi erityisesti sellaisissa kohteissa, joissa toimitaan korkealla tai muuten vaikeapääsyisissä paikoissa. Jotta ihminen pääsisi tällaisiin paikkoihin, pitäisi usein pystyttää kalliita tukirakenteita. Aika, riskit ja kustannukset, jotka liittyvät perinteisiin ratkaisuihin, kun otetaan käyttöön uutta infrastruktuuria tai käytetään raskaita ajoneuvoja, kannustavat kehittämään lentäviä robotteja, joilla pystytään nopeasti saavuttamaan eri työskentelykohteet ja suorittamaan näköhavaintoon tai kosketukseen perustuvia tarkastuksia.

Esimerkkinä hankalia ja vaarallisia tilanteita sisältävästä tehtävästä on nelivuotinen *HYFLIERS-hanke* (*hybrid flying-rolling with snake arm robot for contact inspection*). Siinä kehitetään edistyksellisiä robotteja teollisuuslaitosten tarpeisiin. Tavoitteena on kehittää maailman ensimmäinen teollisuuteen integroitu sekä maassa että ilmassa liikkuva robotti kohteisiin, joissa työskentely on ihmisille vaarallista ja joihin aiemmin kehitetyillä roboteilla ei päästä.

Hanketta viedään eteenpäin yrityskumppanien kanssa. Öljyn- ja kaasunjalostuslaitoksissa hyödynnetään dataa, robotiikkaa ja tekoälymenetelmiä, kun valvotaan laitosten turvallisuutta ja korjaustoimenpiteiden tarvetta tarkastamalla järjestelmän putkia ultraäänellä. Perinteisesti tarkastukset on tehnyt joku työntekijä ja jopa 90-prosenttisesti tikkaiden, telineiden, köysien tai nostureiden avulla. Työntekijät voivat altistua myös korkeille lämpötiloille tai myrkyllisille aineille. Turvatoimista huolimatta onnettomuusriski on korkea.

Testauskohteeksi valittiin öljy- ja kaasulaitokset, koska niissä korroosio ohentaa ajan myötä putken seinämää, mikä voi lopulta hajottaa koko putken. Rikkomattomat mittaukset ovat suositeltavin menetelmä tuotantojärjestelmissä. Normaalikokoisen öljynjalostamon täydellinen tarkastus vaatii noin 50 000 mittausta. Tarkastukset tehdään käsityönä, jolloin henkilökunnan tulee päästä hankaliin ja vaarallisiinkin paikkoihin, mikä voi johtaa putoamisiin tai jopa vammautumisiin. Lisäksi tarkastukset aiheuttavat isoja kustannuksia. Mittauspaikkoihin pääsee usein vain tikkaiden, telineiden, nostureiden tai köysien avulla. Mittauspaikat voivat myös sijaita vaarallisilla alueilla, missä on korkea lämpötila tai myrkyllisiä aineita. Kun henkilökunta joutuu työskentelemään korkeissa ja vaarallisissa paikoissa, vaaditaan etukäteissuunnittelua ja tukihenkilöiden läsnäoloa. Valmistelutyöt maksavat yleensä paljon enemmän kuin itse tarkastusoperaatio.

4.2 Viestit päättäjille ja kansalaisille

Kaksi viimeistä vuotta ovat pandemian vuoksi muuttaneet nopeasti meidän käyttäytymistämme. Opimme yhdessä yössä työskentelemään etänä. Se on tuonut muutoksia myös muihin päivittäisiin rutiineihimme. Verkkokauppa on tullut uusille alueille eikä vähiten elintarvikepuolelle. Tämä tarkoittaa, että nyt rakennetaan entistä parempia, tehokkaampia ja automaattisempia logistiikkajärjestelmiä.

Logististen järjestelmien voidaan odottaa laajentuvan entisestään ja jatkavan kasvuaan. Tämä tarkoittaa panostuksia sekä itseohjautuviin siirtovaunuihin että vapaasti liikkuviin autonomisiin robotteihin (AGV, AMR).

Robotit ovat tulleet palveluroboteiksi myös terveysteknologian alalle. Lääketieteellisten robottien merkitys kasvaa. Työvoimapula ajaa länsimaita kehittämään myös eksoottisempia ratkaisuja, jopa sosiaalisia robotteja. Vuoden 2020 aikana kehitettiin yli 30 desinfiointirobottimallia.

Teollisuusrobotiikassa on nähtävissä tehdassalin ulkopuolella myös raskaiden rakennuskoneiden robotisointi. Kaivospuolelta löytyy tähän jo ratkaisuja pelkästään turvallisuuden vuoksi. Nyt haetaan entistä enemmän mahdollisuuksia käyttää useampaa konetta yhtä aikaa niin, että niitä yksittäinen henkilö. voi huolehtia niiden toiminnasta.

Robottien käyttöön vaikuttavat erilaisten pilviteknologioiden ja tiedonsiirtoteknologioiden kehitys, erityisesti (5G ja sen jälkeiset teknologiat). Reaaliaikainen ja aina luotettava kommunikointinopeus avaa mahdollisuuksia uusille sovelluksille, kuten kirurgisille roboteille. Isot kansainväliset yhtiöt mahdollistavat myös uudenlaisen liiketoimintamallin. Robotti voidaan ajatella palveluna (robot as a service), mikä alentaa kynnystä siirtyä robotiikkaan ja uusille markkinoille.

Perinteistä robotiikkaa täydentävät ihmisen ja robotin yhteistyösovellukset. Päättäjien, kehittäjien ja käyttäjien täytyy ymmärtää selkeästi tekoälyn rajat. Roboteista täytyy tehdä työkavereina läpinäkyviä, niin että ymmärrämme niiden toimintatavat.

Robotiikan pitkän aikavälin näkymät ovat edelleen erinomaiset. Meillä on vieläkin tylsiä, likaisia, herkkiä ja vaarallisia tehtäviä, jotka voidaan tehdä paremmin roboteilla. Näiden työtehtävien robotisointi parantaisi työntekijöiden terveyttä, turvallisuutta ja työtyytyväisyyttä. Ikääntyvissä yhteiskunnissa on yhä enemmän tarpeen vapauttaa työntekijät fyysisistä tehtävistä.

Robotit eivät valloita maailmaa, mutta niitä käytetään yhä enemmän mitä mielikuvituksellisimmissa tehtävissä. Robotiikka tarvitsee tekoälyä ja hyötyy yhteistyöstä tekoälyn kanssa. Tekoälyä ei ole olemassa ilman dataa. Näiden yhteistyöllä on mahdollista synnyttää uusia laitteita, järjestelmiä ja palveluja.

5 Tekoäly lukee, kirjoittaa ja keskustelee

Tapio Salakoski
professori, dekaani
Turun yliopisto

Ihmisen kieli on ihmisen mieli; kieli on korkeamman ajatustoiminnan ja kommunikoinnin luonnollinen väline. Rakenteinen, numeerisesta tai valmiiksi luokitellusta tiedosta koostuva data ei riitä, vaan tarvitaan myös kuvaa, ääntä – ja vapaata tekstiä. Jotta faktat saavat merkityksensä, ne on asetettava kontekstiin: tarina on välttämätön ymmärtämiselle ja oppimiselle. Ihmisen kieli puhutussa muodossa syntyi ryhmän sisäisen kommunikaation tarpeeseen. Tarve kielellisen informaation pysyvään tallentamiseen synnytti sittemmin kirjoitetun kielen, kirjoitusasun lukuisille puhutuille kielille. Kirjapainotaito aiheutti kulttuurisen vallankumouksen ja synnytti modernin sivistyksen.

Koneen kykyä ymmärtää ja tuottaa luonnollista kieltä pidetään koneen älykkyyden olennaisena osana. Sitä tutkiva tekoälyn haara on kieliteknologia eli luonnollisen kielen käsittely (engl. *Natural Language Processing, NLP)*. 1900-luvun puolivälissä kielen kääntämisessä odotettiin tekoälyn nopeata läpimurtoa. Tehtävä osoittautui paljon luultua vaikeammaksi, ja luonnollista kieltä alettiin pitää voittamattomana haasteena tietokoneelle. Nyt kuitenkin kieliteknologian läpimurron aika on käsillä. Viime vuosien nopean kehityksen takana on tietokoneiden laskentatehon ja datan määrän kasvu sekä paradigman muutos. Perinteisestä lingvistisen mallintamisen lähestymistavasta on siirrytty koneoppimiseen ja massiiviseen dataan perustuvaan kieliteknologiaan.

5.1 Mistä on kyse?

Kieliteknologiaa kehitetään kielen käytön eri osa-alueilla ja sovelletaan useisiin tehtäviin. Seuraava jäsennys ei ole tyhjentävä, mutta se antaa kuvan kielen keskeisestä merkityksestä dokumentoidun tiedon kentässä ja kieliteknologian monipuolisuudesta.

5.1.1 Puhe ja teksti

Tämä kirjoitus keskittyy kirjoitettuun kieleen. Puhuttua kieltä voidaan pitää kielen alkuperäisenä muotona, mutta sivistyksen perustana on kirjallinen dokumentaatio. Puheen käsittely on oma tutkimushaaransa, mutta sen tietosisällön käsittely palautuu kirjoitettuun kieleen. Puhe välittää sanallisen sisällön ohella muutakin kommunikaatiota sekä äänensävyn, painotusten ja rytmityksen sisältämää informaatiota, kuten tunteita. Puhuessa tätä informaatiota välittävät myös kasvojen ilmeet ja kehon kieli. Kehon kieltä käytetään myös sanallisen sisällön välittämiseen viittomakielissä.

Puhuttu kieli on myös aina sidoksissa puhujaan: se ei ole anonyymiä, vaan puhesignaalissa kuuluvat puhujan biometriset ominaisuudet. Puhetta voidaan jopa käyttää luotettavasti puhujan tunnistukseen samaan tapaan kuin sormenjälkiä, silmiä tai kasvoja.

Puhetta käsitellään tulkitsemalla vastaanotettu akustinen puhesignaali sanalliseksi informaatioksi tekstittämällä (engl. *Speech to Text*). Sanallinen tietosisältö analysoidaan tekstimuotoiselle datalle kehitetyillä menetelmillä. Puheen tuottaminen puolestaan tapahtuu tekstimuotoisen kielen generoinnin jälkeisellä puhesynteesillä (engl. *Text to Speech*). Sovelluskohteita on lukuisia, ja jopa puhuttu keskustelu on mahdollista.

5.1.2 Kieli ja kielimallit

Moderni datalähtöinen NLP perustuu koneoppimiseen. Uudet tehokkaat koneoppimisalgoritmit vaativat toimiakseen hyvin suuria datamääriä. Tietokone oppii kieltä kuin ihmislapsi: esimerkkien avulla ja toistojen kautta. Puhutut sanat opitaan tunnistamaan ja lausumaan harjoittelun avulla, ja sanojen merkitys täsmentyy niiden esiintyessä yhä uudestaan erilaisissa konteksteissa. Mekanismi on kuin ihmisellä: lukeminen parantaa kielen ymmärrystä, laajentaa sanavarastoa ja vahvistaa ilmaisukykyä.

Kieli voidaan jakaa kielioppiin eli syntaksiin, merkitykseen eli semantiikkaan ja tarkoitukseen eli pragmatiikkaan. Syntaksi kuvaa miten lauseet muodostetaan, semantiikka kertoo mitä niissä sanotaan ja pragmatiikka pyrkii ymmärtämään, miksi ne on kirjoitettu.

Lauseiden rakenteen ymmärtäminen ja sanojen tuntemus ovat välttämättömiä merkityksen ymmärtämiselle. Nykyään kone hallitsee syntaksin hyvin ja kykenee yhä useammin hyödyllisen tasoiseen semanttiseen ymmärrykseen. Luonnollinen kieli on kuitenkin vivahteikasta ja monimerkityksistä. Siksi kielenkäyttäjän todellisen tarkoituksen ymmärtäminen on ihmisellekin haastavaa ja pragmatiikka tekoälylle useimmiten vielä liian vaikeata.

Kirjoitetuilla kielillä on erityispiirteensä, mutta niillä on myös yhteisiä ominaisuuksia. Uusimpien tekoälymenetelmien kyvykkyys perustuu massiivisesta datasta opittuihin kielimalleihin. Ne kiteyttävät kielellisen osaamisen, joka voidaan rajoitetusti siirtää sovelluksesta ja jopa kielestä toiseen. Kielen kääntäminen oli ennen kokonaan ihmisen erikoisosaamista, mutta nykyään koneen tekemä raakakäännös toimii hyvänä pohjana. Vaaditun datan ja laskennan suuren määrän vuoksi kielimallien tuottaminen maksaa paljon, joten ne ovat myös kaupallisesti arvokkaita. Niiden hyvyys riippuu opetusdatan määrästä ja laadusta. Algoritmien monimutkaisuuden vuoksi mallit ovat heikosti selitettävissä: kielimalli on ”musta laatikko”, jota voidaan arvioida vain tulosten perusteella. Tulosten luotettavuus palautuu vahvasti opetusdataan: määrän kasvattaminen parantaa tuloksia. Myös datan laadulla on väliä. Tekoäly oppii esimerkkidatasta samanarvoisesti kaikki sen piirteet, ei pelkästään niitä, joita sovelluksen suunnittelijat tavoittelevat. Jos ja kun opetuksessa käytettävät tekstit sisältävät syrjiviä piilomerkityksiä, ristiriitoja ja epätotuuksia, tekoäly oppii nekin.

5.2 Kieliteknologian käyttötapoja

5.2.1 Tiedonhaku

Erilaiset hakutoiminnot ovat yleisimpiä NLP:n käyttötapoja. Avainsanahausta ollaan siirtymässä kohti vapaan tekstin käyttöä ja monikielisyyttä. Tekstejä voidaan luokitella ennalta annettuihin luokkiin sisällön mukaan tai ryhmitellä niiden samankaltaisuuksien

perusteella. Yksityisten ihmisten suorittamat haut ovat niin yleisiä, että niitä voidaan itsessään käyttää datana esimerkiksi epidemioiden leviämistä selvitettäessä.

Muun muassa terveydenhuollon sovelluksissa hakua usein tehostetaan rakenteistamalla tekstiä etukäteen annetuin avainsanoin tai otsikoin. Muuttuvassa ja epävarmassa maailmassa luotettavimman ja tarkimman tuloksen tuottaa kuitenkin vapaa tekstihaku. Alkuperäinen teksti voidaan jäsentää tai otsikoida myöhemmin uudestaan halutulla tavalla, ja siksi se tulisikin aina säilyttää rakenteistetun dokumentaation rinnalla.

5.2.2 Sisältöanalyysi

Tekstin sisällöllinen analysointi perustuu samankaltaisuuksien, eroavaisuuksien, riippuvuuksien, trendien ja muiden heikkojen signaalien tunnistamiseen. Menetelmiä käytetään diagnostiikan tukena terveydenhuollossa, mutta myös kuluttajien käyttäytymisen selvittämiseksi ja siihen vaikuttamiseksi mainosrahoitteisen toiminnan ja politiikan piirissä. On tärkeätä ymmärtää, että esimerkiksi Facebookin käyttäjät eivät ole palvelun asiakkaita, vaan myytävä tuote. Asiakkaita ovat mainostajat ja muut tahot, jotka ostavat käyttäjäprofiileja toimintansa kohdentamiseksi ja tehostamiseksi. Tällaista toimintaa nähtiin Yhdysvaltain presidentinvaaleissa ja Britannian brexit-äänestyksessä.

Tekstin tiivistäminen ja yhteenvedot, tuotteiden ja palvelujen suosittelujärjestelmät sekä päätöksenteon tuki ovat esimerkkejä sovelluksista, jotka vaativat tekstin syvempää sisällöllistä ymmärtämistä. Laadultaan parantunut konekääntäminen on tulossa reaaliaikaiseen kuluttajakäyttöön. Kasvavia sovellusaloja ovat terveydenhuolto ja kauppa, vakuutus- ja rahoitusala sekä viranomaispalvelut, kuten patentointi ja verotus.

Niin yritykset kuin yhteiskunnalliset tai valtiolliset toimijat ovat kiinnostuneita tekstin aiheen tai aihepiirin, sentimentin (tunnetilan) ja vaikutuksen kohteen tunnistamisesta. Näillä tahoilla on tarvetta tunnistaa vihapuhe, seurata yrityskuvan kehitystä, arvioida poliittisen aloitteen suosiota tai ennakoida päätöksen vaikutusten kohdentumista. Myös kirjoittajan profilointi tai jopa tunnistaminen voi tulla kyseeseen esimerkiksi historiantutkimuksessa, rikostutkinnassa tai turvallisuusvalvonnassa.

Menetelmien käytön sääntely on erityisen tärkeätä silloin, kun käyttökohteet ovat arkaluonteisia ja yksityisyydensuoja voi vaarantua. Mallien monimutkaisuuden vuoksi niiden tarkka kontrolli on kuitenkin vaikeata. Esimerkiksi Google pyrkii sensuroimaan hakutuloksistaan sisältöä, joka mielletään loukkaavaksi. Samalla se kuitenkin tuottaa systemaattisen vääristymän, minkä vuoksi etniset ja seksuaalivähemmistöt ovat tuloksissa aliedustettuina.

5.2.3 Tekstin tuottaminen ja vaikuttaminen

Kielen tuottaminen on luovaa toimintaa, ja puheen tai tekstin tuottaminen tekoälyn avulla vaatii tekoluovuutta. Automaattisen kielengeneroinnin avulla kyetään nopeuttamaan ja laajentamaan tiedonjakoa, tehostamaan viestin perillemenoa, korjaamaan puhujan omia vajavaisuuksia tai tukemaan kirjailijan luovuutta. Puheen tai tekstin piirteitä halutusti muokkaamalla sen avulla voi tehdä myös vahingollisia asioita, kuten matkia muita, loukata immateriaalioikeuksia tai tehdä petoksia. Koska tekoälyllä ei ole

omaa tahtoa, pidemmän ja johdonmukaisesti etenevän tekstin tuottamiseksi tekoäly vaatii vielä ihmisen ohjausta.

Tekoälyllä generoitua kieltä voidaan käyttää henkilökohtaiseen opastukseen ja joukkojen ohjaukseen, mutta myös kaupalliseen tai poliittiseen manipulointiin ja provosointiin. Open AI:n GPT-3-kielimallia käytetään jo yli 300 sovelluksessa, ja se tuottaa noin viisi miljardia sanaa päivässä. Tällaisen kyvykkyyden valjastaminen trollaamiseen voi vaikuttaa vahvasti yleiseen mielipiteeseen. Se voi tukkeuttaa sosiaalisen median ja estää muiden viestien, kuten viranomaisviestien, perillemenon. Kyseessä on selkeästi kansallisen informaatiopuolustuksen piiriin kuuluva uhka, jota vastaan on syytä varautua huolellisesti yhteiskunnan resilienssin turvaamiseksi.

5.2.4 Tukiäly ja tietoisuus

Tekstianalyysi ja tekstin tuottaminen voidaan yhdistää keskustelevaksi tukiälyksi, joka voi sparrata opiskelijoita ja luovan työn tekijöitä, kuten kirjailijoita sekä markkina-analyytikkoja, tuotekehittäjiä ja päätöksentekijöitä. Tällaisia lisätyn luovuuden työkaluja on jo markkinoilla. Ne eivät suinkaan korvaa ihmisen omaa luovuutta eivätkä poista päätöksentekijän vastuuta. Tutkimusten mukaan kaksi kolmasosaa tekoälyn antamista vastauksista sisältyy suoraan opetusdataan; ne on siis opittu ulkoa. Mallit jäljittelevät ihmisiä kaikessa, myös ennakkoluuloissa ja epätotuuksissa, jolloin datan laadun merkitys korostuu. Koska tekoälyltä puuttuu yleinen arkijärki, massiiviset tekoälymallit ovat ihmisiä useammin epätotuudellisia. Rajatuissa aihepiireissä parhaaseen tulokseen päästäänkin pienemmillä räätälöidyillä malleilla, joiden selitettävyys on parempi.

Yleisen tekoälyn edellytyksenä ajatellaan olevan tietoisuuden kehittäminen tai sen kehittyminen itsestään. Ilmiö tunnetaan singulariteettina. Periaatteessa ihmisen kyvyt ylittävän itsetietoisen tekoälyn kehittyminen on mahdollista, mutta yleisen käsityksen mukaan niin kaukana tulevaisuudessa, ettei siitä kannata huolestua. Todelliset uhat tulevat siitä, että ihminen käyttää tekoälyä epäeettisiin ja vahingollisiin tarkoituksiin. Nämä uhat ovat jo olemassa. Vaikka paniikkiin ei olekaan aihetta, on syytä varustautua tulevaisuuteen erityisesti tutkimuksen ja koulutuksen keinoin.

5.3 Tekoäly muutosvoimana

Kieliteknologiaan perustuvat tekoälyratkaisut parantavat yleisesti tiedon saavutettavuutta, ymmärrettävyyttä ja hyödynnettävyyttä. Arvovapaa teknologia on vahvasti mahdollistava niin yksilön, yritysten, yhteiskunnan kuin ympäristönkin näkökulmasta.

5.3.1 Pienten kielten asema

Kieliteknologia madaltaa oleellisesti kielimuuria ja parantaa sitä kautta kansallista ja kansainvälistä vuorovaikutusta, yhdenvertaisuutta ja yhteisöllisyyttä. Toisaalta teknologisen kehityksen yksipuolistuminen uhkaa kaventaa kielellis-kulttuurista diversiteettiä. Samalla on vaarana, että luovuus tukahdutetaan ja inhimillistä innovaatiopotentiaalia heikennetään. Pienten kielten asemasta on kannettava huolta.

5.3.2 Digitalisaation tehostuminen

Kun tietoon pohjautuva päätöksenteko tehostuu, terveydenhuollon, oppimisen ja yhteiskunnallisten palvelujen laatu, saatavuus ja tehokkuus paranevat merkittävästi. Sovellusten yksilöllistyminen parantaa niiden käytettävyyttä ja tarkkuutta sekä edistää yhdenvertaisuutta. Tekoäly voi tukea kulttuuria, luovuutta, luontoarvoja, innovointia, kaupallista toimeliaisuutta ja yrittäjyyttä sekä jopa hidastaa kaupungistumiskehitystä.

Kaupallinen ja yhteisöllinen vaikuttaminen tehostuvat. Fragmentoituva markkina antaa tilaa yhä yksilöllisemmille palveluille ja pienemmille toimijoille. Toisaalta globaali alustatalous uhkaa monopolisoida sovellusten tuottamat taloudelliset hyödyt ja muodostaa merkittäviä ylikansallisia yhteiskunnallisia toimijoita valtioiden rinnalle.

5.3.3 Yksityisyys, väärinkäytökset, valvonta ja vaikuttaminen

Dataistuminen ja digitaalisten palvelujen yhdentyminen uhkaavat yksityisyyttä. Väärinkäytökset yleistyvät, monipuolistuvat ja tulevat vaikeammiksi havaita. Tekoäly mahdollistaa valvonnan merkittävän lisääntymisen, mutta toisaalta se tuo myös uusia ratkaisuja yksityisyyden, luottamuksen ja turvallisuuden parantamiseksi. Yritysten ja yhteiskunnallisten toimijoiden lisäksi myös yksityiset ihmiset alkavat käyttää tekoälyä vaikeasti ennakoitavin tavoin. Tämä luo tarvetta uudenlaiselle sääntelylle.

Yhteiskunnallinen ja kaupallinen vaikuttaminen sekoittuvat ja tehostuvat merkittävästi. Koska teknologiset resurssit jakautuvat epätasaisesti, yhdenvertaisuus saattaa vaarantua. Toisaalta jos fyysiset rajoitteet poistuvat, taloudellinen riippuvuus vähenee ja myös vähemmistöt saavat äänensä paremmin kuuluviin. Myös valtiollinen vaikuttaminen lisääntyy ja kasvattaa merkittävästi kyber- ja informaatiopuolustuksen tarvetta. Yhteiskunnan resilienssi riippuu yhä enemmän kansalaisten yleisestä tieto- ja taitotasosta, yhteisöllisyyden tunteesta ja puolustustahdosta.

5.4 Viestit päättäjille ja kansalaisille

5.4.1 Huippuosaaminen ja yleissivistys

Tutkimus ja koulutus ovat avainkeinoja tekoälyvallankumouksessa pärjäämiseksi. Huippuosaamista tarvitaan kilpailukyvyn parantamiseksi, mutta myös kansallisen valinnanvapauden ja vaikutusmahdollisuuksien varmistamiseksi. Korkea yleissivistys parantaa elämänlaatua ja yhdenvertaisuutta sekä edesauttaa terveiden markkinoiden kehitystä, yhteiskuntarauhaa ja turvallisuutta. Monipuolista kielten osaamista, kulttuurista diversiteettiä ja pieniä kielialueita on tuettava kansallisesti, EU-tasolla ja kansainvälisesti.

5.4.2 Sääntely ja kilpailukyky

Tietosuojan, yksityisyyden, yhdenvertaisuuden ja tasa-arvon sekä yleisen luottamuksen turvaamiseksi tekoälyn käyttöä on säänneltävä kansainvälisesti. Jotta kilpailukykyä voidaan tukea, on pyrittävä määrittelemään tekoäly teknologiariippumattomasti ja asetettava vaatimuksia käytetyn datan laadulle ja eettisesti hyväksyttävälle, kestävän kehityksen mukaiselle käyttötarkoitukselle. Teknologisten ratkaisujen yhteensopivuutta

tulee edistää standardein ja avoimin rajapinnoin. Päätöksenteon vastuut on määritettävä etukäteen, ja datan luovutuksessa informoidun suostumuksen käytänteitä on kehitettävä käyttäjälähtöisesti. Sääntelyssä pitää ottaa huomioon kaupallisten ja yhteiskunnallisten toimijoiden lisäksi yksilöiden aktiivisemmiksi muuttuvat roolit.

5.4.3 Kokonaisturvallisuus

Tietoturva, tietosuoja, yksityisyys ja luottamus ovat avaintekijöitä myös kansainvälisessä vuorovaikutuksessa. Informaatiovaikuttaminen lisääntyy ja monipuolistuu, minkä seurauksena sisäiset ja ulkoiset turvallisuusriskit kasvavat. Näiden turvallisuusriskien ennakointi ja informaatiopuolustuskyvyn parantaminen ovat kansallisen turvallisuuden keskeisimpiä asioita.

5.5 Automatisoituvat kielipalvelut

LINGSOFT on tuonut 35 vuotta kieliteknologiaa jokapäiväiseen käyttöön. Se on ollut mukana kääntämisen ja kielenhallinnan kehityksessä teknologia-avusteiseksi liiketoiminnaksi. Noin sata vakituista työntekijää työskentelee Turussa, Helsingissä ja Tukholmassa satojen alihankkijoiden kanssa. Uuden teknologian avulla Lingsoft on saavuttanut ja pitänyt asemansa Suomen eturivissä ja maailman sadan suurimman kielipalvelujen tarjoajan joukossa. Microsoft, Yle, SVT, eduskunta, Euroopan komissio ja tuhannet muut asiakkaat hyödyntävät Lingsoftin innovatiivisia kieliteknologiaratkaisuja.

Lingsoftin ydinpalvelut ovat kääntäminen ja tekstittäminen. Konekäännös tai puheentunnistus ihmisen tekemin jälkikäsittelyin takaa nykyään parhaan lopputuloksen. Kielityö automatisoituu nopeasti. Yhä enemmän käännöksiä ja tekstityksia tehdään täysin ilman ihmisen apua. Siirtymä automaattisiin ratkaisuihin voi käydä äkillisesti. Esimerkiksi HUS:ssa sadat kirjoittajat irtisanottiin vuonna 2020. Euroopan unioni puolestaan julkaisee jo täysin koneellisesti käännettyjä sivuja. Ratkaisut mahdollistavat suurempien määrien kääntämisen ja tehokkaamman tekstityksen, jolloin tieto voidaan julkaista välittömästi useilla kielillä.

Globaalit tekoälyn yritysjättiläiset hallitsevat julkista keskustelua. Ne eivät priorisoi kaikkia kieliä yhtä lailla, joten on välttämätöntä kehittää paikallisia ratkaisuja pienemmille kielille. Keskeistä on pyrkiä standardointiin ja yhteentoimivuuteen, jotta pienemmät toimijat eivät putoa pois. Muuten suomelle ja ruotsille jäisi vain toisen luokan ratkaisuja. Tämä useita eurooppalaisia kieliä koskeva uhka tunnustetaan jo EU:n tasolla.

5.6 Kilpailukykyä kieliälyllä

Suomalainen **SILO AI** on Euroopan suurimpia yksityisiä tekoälylaboratorioita. Asiakaskunnan ytimessä ovat teollisuus, älylaitteiden valmistajat sekä älykkäät kaupungit. Silo AI auttaa yrityksiä vahvistamaan tuotteidensa kilpailukykyä tekoälyratkaisuin. Silo AI tekee myös omaa tuotekehitystä ja rakentaa kielikeknologiaratkaisuja yhteistyössä merkittävien kansainvälisten toimijoiden kanssa. Niitä ovat esimerkiksi Yhdistyneiden Kansakuntien AIDS:n vastainen UNAIDS-organisaatio sekä saksalaisen vakuutusalan toimijan Allianzin tytäryhtiö IDS Analysis and Reporting Services.

Silo AI:n ratkaisut mahdollistavat rakenteettomien, vapaata tekstiä sisältävien asiakirjojen automaattisen käsittelyn. Valtavasta, vaikeasti käsiteltävästä tietomäärästä voidaan muun muassa analysoida viestien tunnelatausta, arvioida dokumenttien sisältämiä väitteitä sekä tunnistaa ja tarkistaa niissä esiintyviä määreitä, kuten lukuja, nimiä ja aikamääreitä. Kieliteknologian avulla tekstimassoista voidaan tehdä monipuolisesti hyödynnettäviä tietolähteitä. Ne vastaavat niin yritysten, julkishallinnon kuin yksityisten ihmistenkin tietotarpeisiin sekä tukevat päätöksentekoa ja tehostavat asiointia.

Vuodesta 2017 Silo AI on kasvanut ainutlaatuiseksi tekoälyn osaamiskeskittymäksi, jonka pääosaamisalueita ovat kieliteknologia, koneoppiminen ja konenäkö. Silo AI:lla on yli 200 työntekijää, joista tohtoreita on 90. Kieliteknologiaan keskittyy 70 datatieteilijää. Silo AI tähtää Euroopan tekoälykentän lippulaivaksi. Yritys toimii jo Helsingissä, Turussa, Oulussa, Tampereella, Jyväskylässä, Tukholmassa ja Kööpenhaminassa.

6 Tekoäly oppii

Arno Solin
apulaisprofessori, akatemiatutkija
Aalto-yliopisto

Ihmiset ja eläimet omaksuvat asioita havainnoimalla ympäristöään, päättelemällä yrityksen ja erehdyksen kautta tai yleistämällä jo oppimiaan asioita. Tekoälyt lainaavat meiltä tapoja oppia. Yhtäläisyyksiä ihmisen ja tekoälyn oppimisen välillä on kuitenkin vähemmän kuin äkkiseltään luulisi. Tekoälyt kykenevät kuitenkin mukautumaan havaintoihin, oppimaan kokemuksistaan ja selittämään ympäröivää maailmaa.

6.1 Mistä on kyse?

Sana *tekoäly* viittaa moniin erilaisiin menetelmiin ja eri osista koostuviin järjestelmiin. Jo tekoälyn määrittelyn hankaluus antaa vihjeen siitä, että tekoälyt eivät opi yhdellä sel keästi määritellyllä tavalla. Tapa omaksua tietoa riippuu menetelmien rakenteesta ja sovellusalueesta. Kaikki tekoälyt eivät opi, vaan niiden toiminta voi myös perustua ihmisen määrittelemään käytökseen. Tällaisia ovat esimerkiksi jotkin pelitekoälyt, jotka seuraavat niille syötettyjä lainalaisuuksia (esim. pelin säännöt), ja älykäs termostaatti, joka seuraa lämpötilaa ja säätelee lämmitystä annettujen havaintojen perusteella. Moni nykyisistä tekoälyjärjestelmistä perustuu kuitenkin koneoppimiseen eli tekoälyn kouluttamiseen havaintoaineiston eli datan avulla.

Koneoppiminen (engl. *machine learning*) viittaa järjestelmiin, jotka oppivat havainnoimalla tietoaineistoja ja kykenevät seuraamaan löytämiään lainalaisuuksia. Koneoppimisessa ei tyypillisesti tavoitella ihmisen kaltaista yleistä älykkyyttä. Sen sijaan siinä koulutetaan menetelmiä, jotka kykenevät tekemään yksittäisiä samankaltaisia asioita, joita me älykkäät olennot kykenemme tekemään.

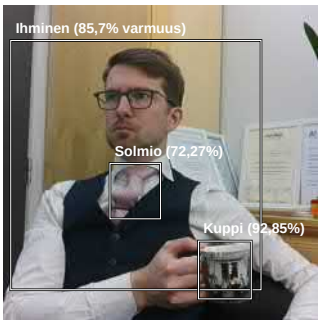
Kuvitellaan kaksi ääripäätä. Toisessa ääripäässä mallinnettavasta ilmiöstä tiedetään kaikki eikä havaintoja tarvita (esim. shakin pelaamisessa säännöt). Toisessa ääripäässä sen sijaan ei tiedetä ennalta mitään mutta kaikki voidaan havainnoida tai mitata (esim. kaikkien kansalaisten syntymäajat). Kummassakaan ääripäässä koneoppimista ei tarvita, sillä kaikki mahdollinen joko tiedetään tai havainnoidaan. Koneoppimista tarvitaan näiden kahden ääripään välissä, jossa on tarpeen yhdistää ennakkotietoa ja havaintoja opittavasta ilmiöstä. Ennakkotieto voi koostua myös oletuksista. Oletusten ja havainnoinnin tasapaino vaihtelee sovelluksen mukaan. Fysiikan lait on helppo kertoa tekoälylle matemaattisessa muodossa, kun taas kissan kuvailu kaavoilla on vaikeaa. Tämän takia jotkin ihmiselle helpot asiat ovat tekoälylle vaikeasti opittavia – ja toisin päin.

Koneoppiminen on käytännössä kolmen osatekijän summa: mallin, havaintoaineiston ja koulutusmenetelmän. Ensinnä tarvitaan matemaattinen **malli**, joka sisältää mallioletukset ja jonka rakenne pitkälti määrää, mitä tekoäly kykenee oppimaan. Käytännössä malli sisältää vapaita parametreja. Mitä enemmän on vapaita parametreja, sitä joustavampi

malli yleensä on ja sitä monimutkaisempiin opittaviin asioihin se kykenee mukautumaan. Mallioletukset voivat olla joko eksplisiittisiä tai implisiittisiä. Eksplisiittiset mallioletukset koostuvat esimerkiksi luonnonlaeista, säännöistä tai toiveista. Implisiittiset mallioletukset johtuvat puolestaan mallin rakenteesta. Toiseksi tarvitaan **havaintoaineisto**, joka kuvaa opittavaa tehtävää. Kolmanneksi tarvitaan **menetelmä** (algoritmi), joka kykenee yhdistämään mallin ja aineiston. Malli koulutetaan selittämään annettu havaintoaineisto. Käytännössä parametreille pyritään siis löytämään arvot, joiden avulla malli pystyy tuottamaan halutun lopputuloksen. Tavoitteena on, että koulutettu malli pystyy tämän jälkeen päättelemään asioita myös sellaisesta datasta, jota se ei nähnyt opetuksen aikana, eli yleistämään oppimiaan asioita.



KUVA 1. Ohjaamattomasti eläinkuvien avulla opeteteun tekoälyn luomia kuvia eläimistä.



Ohjatusti opetetun tekoälyn nimeämiä asioita ja esineitä kuvassa (henkilö, solmio, kuppi).



Vahvistusoppimalla toimiva robotti, joka opettelee kävelemään yrityksen ja erehdyksen kautta.

6.1.1 Ohjaamaton oppiminen

Ohjaamattomassa oppimisessa (engl. *unsupervised learning*) tekoälylle esitetään tietoaineisto sellaisenaan. Koulutusvaiheen tarkoituksena on antaa tekoälyn löytää rakennetta annetusta aineistosta omatoimisesti. Tavoitteena voi olla esimerkiksi kouluttaa tekoäly, joka kykenee luomaan kuvitteellisia ihmisten kasvokuvia. Tyypillinen tapa kouluttaa tällainen malli on esittää tekoälylle tuhansia kuvia ihmiskasvoista ja antaa tekoälyn oppia, mitä sen odotetaan pystyvän tuottamaan. Tällöin tekoälylle ei kerrota kasvojen rakenteesta (esim. missä nenän kuuluu olla tai että silmiä on yleensä kaksi) vaan sen annetaan itse löytää lainalaisuudet ja rakenne aineistosta. Kuvassa 1 vasemmalla on hyödynnetty ohjaamatonta oppimista kuvitteellisten eläinkuvien luomiseen.[1]

Ohjaamatonta oppimista käytetään myös silloin, kun halutaan tarkastella monimutkaista tietoaineistoa ja löytää siitä lainalaisuuksia ja säännönmukaisuuksia. Yksinkertainen esimerkki tästä on vaalikonevastauksien pääkomponenttianalyysi (esim. vaalikonevastauksien visualisointi nelikenttänä sanomalehdissä), joka yleensä löytää liberaali–konservatiivi- ja vasemmisto–oikeisto-akselit vastauksista ilman, että näistä olisi kerrottu menetelmälle etukäteen. Ohjaamaton oppiminen tarjoaa siis työkaluja sekä monimutkaisten

1 Karras, Tero, ym. *"Alias-free generative adversarial networks". Advances in Neural Information Processing Systems* (2021).

asioiden yksinkertaistamiseen (selittäminen) että havaintojen generointiin, muokkaukseen ja tulkintaan.

6.1.2 Ohjattu oppiminen

Ohjatussa oppimisessa (engl. *supervised learning*) tekoäly koulutetaan näyttämällä sille esimerkkejä sekä syötteistä (mitä menee mallin sisään) että toivotuista tuloksista (mitä tulee mallista ulos). Tavoitteena on kouluttaa tekoäly, joka kykenee opetuksen jälkeen yleistämään opittuja asioita ja päättelemään, mitä tuloksia tuntemattomilla syötteillä tulisi saada aikaan. Ohjattu oppiminen on yleisesti käytetty tapa opettaa regressio- ja luokittelumalleja. Regressiomalli pyrkii ennustamaan tai selittämään jatkuvia suureita. Sen avulla voidaan esimerkiksi ennustaa lapsen pituus aikuisena, kun vanhempien pituudet ovat tiedossa. Luokittelumalli pyrkii nimensä mukaan luokittelemaan asioita joko kahteen tai useampaan luokkaan. Sitä voidaan käyttää esimerkiksi tunnistamaan, onko kuvassa kissa vai koira.

Ohjattu oppiminen auttaa tyypillisesti tekoälyä oppimaan sellaisia tehtäviä, joista on kertynyt havaintoaineistoja, joiden lopputulos tiedetään tai joihin on liitetty ihmisen tavoittelema päämäärä. Tällainen tietoaineisto voi olla esimerkiksi kokoelma kuvia esineistä, jotka ovat ihmisen nimeämiä. Näin voidaan kouluttaa malli, joka kykenee tunnistamaan esineitä ja nimeämään ne. Ohjattu oppiminen muodostaa pitkälti monen nykyisen tekoälyjärjestelmän selkärangan.

6.1.3 Vahvistusoppiminen

Vahvistusoppiminen (engl. *reinforcement learning*) sopii käyttökohteeseen, jossa on helppo määritellä päämäärä oppimiselle (esim. pelin voittaminen tai robotin liikkuminen eteenpäin) mutta jota on hankala kuvailla matemaattisessa muodossa tai josta on vaikea kerätä sopiva opetusaineisto. Esimerkki vahvistusoppimisesta on robotin opettaminen kävelemään. On hankala määritellä, miten raajojen nivelten pitää käyttäytyä, jotta tasapaino säilyy ja robotti pääsee samalla eteenpäin – etenkin epätasaisessa maastossa. Sen sijaan on helppo määritellä, milloin oppiminen on johtanut tavoiteltuun lopputulokseen eli on päästy etenemään.

Vahvistusoppimisessa tekoäly tasapainoilee eksploitaation, eli jo opittujen tietojen ja taitojen hyödyntämisen, ja eksploraation, eli uuden tutkimisen ja yrittämisen, välillä. Tähän törmäämme me ihmisetkin arjessa: tilaanko ravintolassa vanhan tutun annoksen (eksploitaatio) vai kokeilenko listalta jotain uutta (eksploraatio)? Tässä on nähtävissä analogia ihmislapsen liikkumiseen. Alussa lapsi ei osaa hallita raajojaan, mutta yrityksen ja erehdyksen kautta hän oppii kääntymään, konttaamaan ja lopulta kävelemään. Vahvistusoppiminen on nouseva tekoälyn osa-alue, jossa läpimurtoja on tapahtunut aivan viime vuosina. Google DeepMindin AlphaZero-tekoäly[2] hyödynsi harjoittelussaan vahvistusoppimista, kun se päihitti ensimmäisenä parhaat ihmispelaajat shakissa, gossa ja shogissa. Se oli opetellut pelistrategioita pelaamalla itseään vastaan.

---

2 Silver, David, ym. *"A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play."* Science 362.6419 (2018): 1140–1144.

6.2 Tekoäly muutosvoimana

Tekoälyt ovat nopeasti siirtyneet akateemisesta tutkimuksesta jokapäiväiseen käyttöömme internetin, mobiililaitteiden ja kuluttajatuotteiden kautta. Oppivien tekoälyjen rivakkaa kehitystä on ajanut kolme tekijää:

- Perustutkimus tekoälymenetelmien kehittämiseksi on mahdollistanut sekä monipuolisemmat mallit että tehokkaammat opetusmenetelmät.
- Tekoälymallien opettamiseen tarvittavan datan määrä on lisääntynyt (esim. avoimet kuva- ja tekstiaineistot).
- Laskentakapasiteetti on kasvanut ja päätelaitteet ovat yleistyneet (esim. älypuhelimet). Niiden tarjoamat ekosysteemit mahdollistavat kaupallisen toiminnan.

Tulevat kehitysaskeleet näillä kaikilla kolmella osa-alueella tukevat uusien tekoälyyn pohjautuvien innovaatioiden ja avausten syntymistä.

Muutos on ollut ripeää ja monet seikat on huomioitu käytännön näkökulmista vähän myöhemmin. Koneoppimismalleissa mahdollisia kompastuskohtia voi olla niin opetusaineistossa, mallioletuksissa kuin mallin opettamisessakin. Pahimmillaan opetusaineisto sisältää sinne kuulumatonta tai ei-toivottua dataa. Tekstiaineisto voi sisältää esimerkiksi vihapuhetta. Aineisto ei myöskään välttämättä ole kattavaa, jos esimerkiksi murresanat puuttuvat tekstiaineistosta. Toisaalta mallioletukset saattavat olla vääriä tai ne eivät huomioi sovelluksen erityispiirteitä. Opetus puolestaan saattaa olla tehotonta tai epäonnistua kokonaan. Näiden seikkojen huomioiminen on korostunut viimeaikaisessa tutkimuksessa samoin kuin tekoälyjen selitettävyyden ja tulkittavuuden parantaminen: mitä tekoäly oppii, tai miten se vetää johtopäätöksensä?

6.3 Huomioitavaa päättäjille ja kansalaisille

Ihmisen oppiminen on jatkuvaa ja jopa tahatonta. Tekoälyt sen sijaan eivät tyypillisesti opi jatkuvasti, vaan niiden kouluttaminen ja käyttäminen ovat erillisiä prosesseja. Jotta tekoäly oppii uutta, se on käytännössä opetettava uudelleen ja otettava uudelleen käyttöön. Tämä tarkoittaa myös sitä, että tekoäly ei tyypillisesti opi omista virheistään. Jos se erehtyy kerran, se erehtyy jatkossakin samassa tilanteessa, jos sitä ei kouluteta uudelleen. Ihminen kykenee huomioimaan epävarmuutta ja tietyissä rajoissa ilmaisemaan oman ymmärryksensä rajat. Tekoälyt eivät pysty hallitsemaan epävarmuuksia luonnostaan, vaan ne saattavat olla ylitsevuotavan varmoja asioista, joista eivät tiedä opetusaineiston perusteella mitään. Esimerkiksi tekoäly, joka on opetettu vain luokittelemaan kissoja ja koiria kuvista, saattaa olla täysin varma, että autoa esittävä kuva on koira. Epävarmuuksien hallinta vaatii erityistä huomioimista mallioletuksissa, opetusvaiheessa ja menetelmien käytössä.

Oppivien tekoälyjen käytettävyyteen vaikuttaa tarjolla oleva opetusaineisto. Tynnyrissä kasvanut tekoäly ei ymmärrä maailmaa. Pienenä kielialueena ja yhteiskuntana, jolla on omat kansalliset erityispiirteensä, meidän pitää tukea ja turvata kotimaisia erityispiirteitä sisältävien opetusaineistojen saatavuus. Tämä koskee kielellisiä aineistoja sekä aineistoja, jotka koskevat esimerkiksi Suomen luontoa, geeni- ja terveystietoja, tieverkoston erityispiirteitä, elinkeinorakennetta tai lainsäädäntöä.

Tekoäly ei itsessään päätä, mitä se oppii, vaan se mukautuu aineiston rakenteeseen ja oppii peilaamaan opetusaineistoaan mallioletustensa antamin reunaehdoin. Tästä syystä on tärkeää, että mallien rakentajat ymmärtävät, mitä ovat tekemässä. Se puolestaan vaatii resursseja koulutukseen.

6.4 Oppivat tekoälyt täydentävät ihmisen osaamista

Havaintoaineistosta oppivat tekoälyt ovat yleisessä käytössä päätöksenteon ja havainnoinnin tukena, data-analytiikkatyökaluina ja tekemässä suosituksia. Koneoppimismenetelmät täydentävät monilla aloilla perinteiseen tilastotieteeseen pohjautuvia analyysimenetelmiä. Ne ovat mahdollistaneet uusia teknisiä avauksia palvelujen ja kuluttaja-tuotteiden osina. Suomalaisessa teollisuudessa monia näistä menetelmistä käytetään täydentämään jo olemassa olevia järjestelmiä ja tehostamaan toimintaa. Samalla kuitenkin kehitysaskeleet esimerkiksi havainnoinnin, kuten konenäön, parissa mahdollistavat uusia avauksia sekä entistä autonomisempia ja automatisoidumpia ratkaisuja. Tällaisia ovat esimerkiksi autonomiset työkoneet, kuljettajaa helpottavat avustimet autoissa, robottivarastot tai tuotantoprosessien automaattinen valvonta. Oppivia tekoälyjä hyödynnetään myös konekääntämisessä (esim. Google-kääntäjä), asiakaspalvelun analysoinnissa (esim. Kansaneläkelaitos), mainosten ja palveluiden kohdentamisessa sekä lääketieteellisissä sovelluksissa.

7 Tekoäly luo

Pertti Saariluoma
kognitiotieteen professori
Jyväskylän yliopisto

7.1 Johdanto

Luovuudella tarkoitetaan ihmisen kykyä keksiä uusia ja tärkeitä asioita. Luovissa prosesseissa ongelmien ratkaisut käsitetään yhtenä hetkenä yhdellä tavalla ja hetkeä myöhemmin ne saavat uuden muodon. Tärkeää luovuudessa on se, että uudet asiat liittyvät olennaisiin kysymyksiin. Täysin vapaat assosiaatiot ja mielleyhtymät ovat harvoin luovia, koska niiden lopputulokset ovat turhia ja pahimmillaan johtavat erilaisiin virheisiin. Luovuus on yhtä hyvin yksittäisen ihmisen, yksittäisen organisaation kuin kansakunnan tai ihmiskunnan ominaisuus. Voidaan puhua psykologisesta ja historiallisesta luovuudesta. Darwin oli yksilönä luova keksiessään ja konstruoidessaan kehitysteorian. Toisaalta kehitysteoria empiirisesti tuhansin kerroin laji-, solu- ja molekyylitasolla vahvistettuna ajattelutapana on saanut vankkumattoman aseman biologian alalla. Käytännöllisten sovellutustensa seurauksena Darwinin ajatus on saanut historiallisen luonteen.

Luovuuden lähtökohtana on ihmisen ajattelu. Ajattelullahan tarkoitetaan prosessia, joka syntyy, kun ihmisillä on päämäärä, mutta he eivät tiedä välittömästi, miten tavoitteen voisi saavuttaa. Tavoitteen saavuttaminen edellyttää tällaisissa tilanteissa uusien ajatusten keksimistä. Esimerkiksi koronaan liittyviä terveysongelmia on voitu ratkaista keksimällä uusia rokotteita ja viemällä niitä tuotantoon.

Ihmisten luovaan ajatteluun liittyy nykyaikana sosiaalinen konstruktiivisuus. Jumbojetissä on satojatuhansia osia. Sellaista ei pysty yksi ihminen luomaan. Samalla tavalla yhteiskunnat tulevat päivä päivältä monimutkaisemmiksi. Sen vuoksi ne eivät ole yksittäisen henkilön rakennettavissa, vaan muurahaispesän tavoin ne koostuvat tuhansien yksilöiden ajattelun ja yhteistyön tuloksista.

Uudet innovatiiviset ideat johtavat yleensä uusiin sosiaalisiin haasteisiin ja ratkaisuihin. Sähkö muutti yhteiskunnan täysin uuden kaltaiseksi. Tänä päivänä suuret uudistukset perustuvat digitalisoinnille ja älykkään teknologian kehittämiselle. Yksi tekoälyn tuoma lisä ihmisten toimintaan on sen kyky luovaan informaation prosessointiin.

Tekoälyllä tarkoitetaan tässä yhteydessä tekniikkaa, joka kykenee suoriutumaan ihmisiltä älykyyttä vaativista tehtävistä. Tällaisia tehtäviä voivat olla esimerkiksi kuljetusten automatisoiminen, älykkäät valmistusprosessit, robottilääketiede tai uudet opetustavat. Tällaiset perinteiset työtehtävät ovat kuuluneet aiemmin vain ihmisille, mutta tekoälyn ansiosta niiden toteuttamiseen on tullut monia uusia tapoja. Nyt ja lähitulevaisuudessa on mahdollista toteuttaa monimutkaisiakin ohjausprosesseja tekoälyn ja muiden älykkäiden teknologioiden avulla.

Tekoälysovellusten avulla toteutettavat informaatioprosessit tuottavat aidosti uusia ja relevantteja ratkaisuja. Esimerkkinä tällaisista voi mainita perustutkimuksena paljon tutkitut ohjelmistot, jotka pelaavat erilaisia pelejä. Jo 1960-luvulla toteutettiin tammea pelaava oppiva ohjelma CHECKERS. Se alkoi aamulla pelaamaan ihmisten kanssa, ja iltapäivällä se oli oppinut voittamaan vastustajansa. 1990-luvulla shakkia pelaavat ohjelmat kykenivät voittamaan jopa maailmanmestarin. Myöhemmin vastaavia ohjelmistoja on toteutettu monissa muissakin toimintaympäristöissä.

Edellä mainituissa esimerkeissä lähinnä ”leikkikaluympäristöissä” toimivat ohjelmistot kykenivät parempaan suoriutumiseen kuin ammatikseen alalla toimivat ihmiset. Esimerkit osoittavat, että monia luovia informaatioprosesseja pystytään toteuttamaan paremmin tekoälyn avulla kuin ihmisten tekeminä. Esimerkiksi suurten tietomassojen käsittelyssä tekoäly avaa monia tärkeitä mahdollisuuksia.

Tässä yhteydessä on hyvä käsitellä kolmea päänäkökulmaa luovaan tekoälyyn. Ensimmäinen niistä koskee sitä, mikä tekee mahdolliseksi tekoälyohjelmistoille olla luovia. Toinen tarkastelukulma koskee sitä, miten ihmiset käyttävät tekoälyä hyväksi luovissa toiminnoissaan. Esimerkiksi muusikot ovat jo vuosia hyödyntäneet tietokoneita säveltämisessä. Kolmantena voidaan kiinnittää huomiota siihen, millaisia uusia mahdollisuuksia tekoäly avaa teknisten, organisatoristen ja yhteiskunnallisten prosessien kehittämisessä.

7.2 Miksi tekoälyjärjestelmät voivat tuottaa luovia tuloksia?

Luova toiminta lähtee luovista ajatuksista. Luovat ajatukset puolestaan ovat todellisuutta kuvaavaa informaatiota ihmisten mielissä eli mentaalisia representaatioita. Näiden voidaan katsoa olevan myös tietynlaisia informaatiotiloja ihmisten mielissä. Esimerkiksi sanomalehden otsikko ”Kaskinen voitti Kuusamon”, jossa paikkakuntien nimet viittaavat teollisuuden sijoituspaikkoihin, esittää yhtä asiantilaa. Se on informaatiota, joka yksittäisen lukijan mielessä se muuttuu ajatukseksi.

Ajatukset ovat informaatiotiloja tai uskomuksia siitä, miten asiat ovat tai miten niiden tulisi maailmassa olla. Tietokoneiden erityisominaisuus on kyky tuottaa nopeasti suuria määriä erilaisia informaatiotiloja. Tämä tapahtuu yleensä joko konstruoimalla uusia ideoita tai ”suodattamalla” jo olevasta tietomassasta sopivia vaihtoehtoja. Usein tietokoneiden tai tekoälyn tuottamat uudet informaatiotilat ovat sisällöltään aiemmin tuntemattomia, ja tässä mielessä tekoälyjärjestelmät voivat olla luovia. Ne voivat esimerkiksi havaita uusia yhteyksiä erilaisten asioiden välillä. Tekoälyjärjestelmät voivat esimerkiksi karsia tarpeettomia sairaalakäyntejä ohjaamalla potilasvirtoja aiempaa tarkoituksenmukaisemmin, helpottaa lääkkeiden kehitystyötä tai edistää terveellisempiä elämäntapoja.

Mikä hyvänsä tekoälyn kehittämä idea ei ole hyvä. Systeemit voivat tuottaa valtavalla nopeudella puhdasta puppua. Systeemi voi tulkita mustia laseja käyttävän nelivuotiaan lapsen filmitähdeksi, epäonnistujaksi tai vaikka gangsteriksi. Siitä huolimatta sotkun seassa voi olla informaatiotiloja, jotka ovat tosia ja olennaisia. Oikeastaan ideoiden kannalta on yhdentekevää, ovatko ne ihmisten tai tekoälyn tuottamia. Olennaista on se, että ne vastaavat sitä, miten niiden esittämät asiat ovat maailmassa ja että on hyviä syitä uskoa, että asiat ovat, kuten idea esittää niiden olevan.

Tekoälyn luovuutta voidaan tarkastella useasta eri näkökulmasta. Yksi tavallinen jako tehdään kombinatorisen luovuuden, tutkiskelevan luovuuden ja muuntavan luovuuden välillä. Ensin mainitulla tarkoitetaan sitä, että tuttuja ajatuksia liitetään toisiinsa erilaisilla uusilla tavoilla. Tätä voidaan tehdä tietokoneilla nopeasti. Siten on mahdollista etsiä erilaisia vaihtoehtoja valtavista vaihtoehtojen puurakenteisista jonoista. Ongelmana tässä ajattelutavassa on se, että mahdollisten yhdistelmien joukot kasvavat nopeasti valtaviksi.

Annetaan kymmenelle apinalle kymmenen kirjoituskonetta ja rajattomasti aikaa takoa uusia kirjainjonoja koneilla. On selvää, että ennemmin tai myöhemmin tätä niin sanottua British Museumin algoritmia käyttävät apinat kykenevät tuottamaan kirjain kirjaimelta Shakespearen kootut teokset. Ongelmaksi muodostuu kuitenkin se, että Shakespearen koottujen teosten tuottaminen tällä tavalla kestäisi valtavan pitkän ajan. Tuona aikana tuotettaisiin määrättömästi merkityksettömiä kirjainjonoja ja vain muutamia merkitseviä representaatioita. Sen lisäksi lukutaidottomat apinat eivät kykenisi tietämään, milloin lopputulos on saavutettu tai mitä tarkoittaa Hamletin kuuluisa lause ”ollako vai eikö olla”. Joten ne jatkaisivat naputtelua huomaamatta löytäneensä ratkaisun. Rajoitetuissa toimintaympäristöissä, kuten shakkipelissä, tällainen sokea etsintä toimii kuitenkin aika hyvin.

Tutkiskeleva luovuus lähtee liikkeelle jostakin olemassa olevasta piirrejoukosta ja muuntelee sitä. On mahdollista antaa tekoälylle syötteenä suuri joukko maalauksia ja sallia sen tuottaa taidesuuntauksen sääntöjä noudattaen uusia maalauksia. Asiantuntijatkaan eivät välttämättä suoralta kädeltä kykene sanomaan, mitkä ovat ihmisen ja mitkä koneen tekemiä maalauksia. Kone ei kuitenkaan kykene luokittelemaan teoksia hyviin ja huonoihin.

Muunteleva luovuus on tutkiskelevan luovuuden kehittyneempi muoto. Muuntelevassa luovuudessa tekoälysysteemi muuntelee itse toimintakehyksen määrittelyä. Se poistaa oletuksia ja lisää niitä luoviin ohjelmistoihin. Tämän seurauksena syntyy aivan uudenlaisia ja ennennäkemättömiä yhdistelmiä. Jos esimerkiksi yritys suunnittelee toimintaansa tiettyjen teknologiaodotusten pohjalta, muuntelevan luovuuden näkökulmasta voi ajatella, mikä olisi mahdollista, kun itse oletuspohjaa muunnellaan. Voitaisiin vaikka kysyä, mitä tapahtuisi, jos materiaalit kestäisivät paremmin kuumuutta, ja etsiä tekoälyn avulla mahdollisia ratkaisuja.

Kaikkia yllä mainittuja peruslähestymistapoja voidaan soveltaa tilanteisiin, joissa tekoälyn luovuutta tarvitaan. Ratkaisemattomaksi ongelmaksi jää kuitenkin aina sen päättäminen, mitkä tekoälyjärjestelmän tuottamista vaihtoehdoista ovat olennaisia ja mitkä ovat roskaa. Koneet eivät tiedä, mikä on olennaista tai relevanttia. Niille kaikki vaihtoehdot ovat samanarvoisia, ellei ohjelmoija ole koodannut ohjelmistoon mahdollisia tuotoksia tai ideoita arvioivia kriteerejä. Relevanssien määritteleminen on ihmisen tehtävä, ja siitä syystä kaikki tekoälyohjelmistot ovat pakosta ihmisen toiminnassa apuvälineitä tai työkaluja.

7.3 Tekoäly avustaa luovia toimintoja

Kun shakin maailmanmestari Garri Kasparov oli hävinnyt ottelunsa tietokoneelle, hän päätti uusia kokeen. Tällä kerralla Kasparov otti mukaansa toisen tietokoneen ja pelasi sen avustamana. Lopputulos oli Kasparoville suosiollinen. Maailmanmestariohjelma hävisi ihmisen ja koneen tiimille. Tämä esimerkki valaisee hyvin sitä, mikä on tekoälyn luovuu-

dessa kriittistä. Luovakin tekoäly on loppujen lopuksi parhaimmillaan ihmisen apuvälineenä. Tällä tavalla symbioosissa ihmisen kanssa se voi edistää merkittävästi ihmisenkin luovuutta.

Tekoälyä käytetään nykyään moniin ihmisten luovuutta edistäviin tehtäviin. Analysoimalla laajoja potilastietokantoja on mahdollista löytää monille sairauksille uusia selittäviä tekijöitä. Sen seurauksena näihin sairauksiin voidaan myös kehittää uusia ja aiempaa parempia hoitomuotoja. Ihmisten kannalta tekoälyjärjestelmät ovat arvokkaita, koska ne voivat tuottaa luovia näkökulmia käsillä oleviin haasteisiin.

Älykkään teknologian ja ihmisen yhteistyö on tärkeä asia pitää mielessä, kun suunnitellaan tekoälysovellutuksia. Amerikkalaisessa, 1500 yritystä käsittäneessä tutkimuksessa osoitettiin, että yritykset kykenivät parhaiten lisäämään omaa suorituskapasiteettiaan, kun ihmiset ja tekoäly paransivat toistensa toimintaa interaktiivisesti. Ihmiset voivat hyötyä tekoälyn tuottamista tiedoista ja ratkaisumalleista, mutta opitun perusteella he voivat myös muokata tekoälytoimintojen parametreja tarkoituksenmukaisempaan suuntaan.

Tekoäly voi toimia yhdessä ihmisen kanssa tiiminä, jossa tekoälyjärjestelmällä voi olla erilaisia rooleja taustadatan tuottamisesta jonkinasteiseen johtamiseen. Oikein ohjattu tekoäly kykenee tarkastamaan paitsi tietoja myös sitä, miten tehtävät on toteutettu. Luova tekoäly voi jopa ehdottaa uusia toimintamalleja.

Ihmisen ja tekoälyn yhteistyö on aina dynaaminen prosessi, jossa kokonaistoimintoa voidaan jatkuvasti kehittää aikaisemmin saatujen kokemusten pohjalta. Luovien prosessien kannalta toimintojen jatkuva kehittäminen onkin avainasemassa. Luovaa tekoälyä käyttävien organisaatioiden on syytä alusta lähtien asennoitua kokeilemaan, luomaan ja testaamaan. Oikea asennoituminen on luovan tekoälyn kannalta olennainen asia.

7.4 Käytännön esimerkkejä

Luovalla tekoälyllä on jo nyt monia toimivia käytännön sovellutuksia. Sekä kuvataiteessa että musiikissa on ollut erilaisia tekoälyyn perustuvia taideteosten tuotantoprosesseja jo vuosien ajan. Esimerkiksi suuri joukko impressionististen taideteosten kuvia voidaan antaa syötetietoina tekoälyohjelmistolle, jonka tehtävänä on yhdistellä niistä uusi työ. Tekoälyn tuottamaa taideteosta on hyvin vaikea erottaa ihmisen tuottamista teoksista. Hieman tavallisempia esimerkkejä ovat monet uudet ruokareseptit. Nekin voivat olla vahvasti tekoälyn tukeman luovuuden tuloksia. Käyttäjä voi antaa koneelle syötteenä mieliruoka-aineitaan, ja algoritmi pohtii, millaisen uuden ruokalajin niistä voisi kehittää.

Autoteollisuus on jo pitkään käyttänyt erityyppisiä älykkäitä komponentteja autoissa. Suurin läpimurto tulee olemaan kuitenkin itseään ohjaavien autojen kehittäminen. Myös erilaiset tekoälyn ohjaamat pakettien kuljetusjärjestelmät ovat tekemässä läpimurtoaan. Jo nyt uutisista voi lukea, kuinka Virossa osa pakettien kuljetuksista nojaa robottikuljettimiin. Aivan vastaavasti sotilastoiminnassa on käytetty jo pitkään risteilyohjuksia, jotka ohjaavat itse itseään ja saattavat vaihtaa maalia itsenäisesti.

Tekoälyyn perustuva valvonta on tavallista monissa maissa, kuten Kiinassa. Kasvojen tunnistusohjelmat ovat aktiivisesti käytössä maassa. Niitä voidaan kuitenkin käyttää hyväksi myös valvottaessa asiattomien kulkua kouluissa tai työpaikoilla. Länsimaissa näihin käyttötapoihin liittyy kuitenkin usein sekä eettisiä että juridisia rajoitteita.

Erilaiset pankkisovellukset ovat myös tavallisia. Tekoälyn avulla voidaan analysoida pörssikursseja tai pyrkiä ennustamaan talouden kulkua. Markkinoinnissa ja politiikassa voidaan hyödyntää neuroverkkoja ja niiden antamia tietoja ihmisten elämästä. On jopa olemassa älykkäitä sovellutuksia, joita käytetään yritysten hallitustyön edistämiseen. Lääkärit voivat käyttää tekoälyä tukemaan riskianalyysejä, mallintamaan molekyylejä tai etsimään syöpiä kuvantamistuloksista. Potilaan antamia tietoja voidaan kerätä puheesta automaattisesti, ja tekoälyohjelmistot voivat analysoida niistä ilmeneviä terveysongelmia Robotteja käytetään kokeellisesti osana leikkauksia, koska tällä tavalla voidaan lisätä terveydenhoidon suorituskapasiteettia merkittävästi.

Tekoälyn avustamat toiminnot ovat levittäytymässä askel askeleelta ja sovellutusalue sovellutusalueelta osaksi koko yhteiskuntaa. Jatkuvasti kehitetään olemassa olevia ratkaisuja ja keksitään uusia, tekoälyyn pohjautuvia tai tekoälyn tukemia toimintoja. Tekoälyn nopeus ja suuri tiedonanalyysikapasiteetti tekevät mahdolliseksi uudistaa yhteiskunnan arkista luovuutta merkittävästi.

7.5 Luova uudistaminen tulevaisuudessa

Luova tekoäly ja tekoälyn käyttö ovat syntyneet jo varsin aikaisessa älykkään teknologian kehittämisvaiheessa 1960-luvulla. Tutkijat ideoivat neuroverkkomalleja jo 1950-luvun lopulla, ja samaan aikaan luotiin perusta symboliselle laskennalle. Tällöin luovan tekoälyn tekniset perusratkaisut saivat muotonsa. Vuosikymmenten kuluessa sekä algoritmit että tietokoneet ovat kokeneet luovuuden näkökulmasta läpimurron toisensa jälkeen.

Ei ole mitään syytä olettaa, etteikö kehitys seuraavina vuosikymmeninä nopeutuisi sitä mukaa kun tietotekniikka kehittyy. Uutena tekijänä on kuitenkin tullut esiin se, että tekoälyyn pohjautuvat toiminnot kokevat myös merkittäviä muutoksia. Hyvä esimerkki on automaatiotyöttömyys. Se tarkoittaa sitä, että älyllisesti yhä vaativammissa tehtävissä tullaan käyttämään älykkäitä järjestelmiä. Sen seurauksena tehtäviä aiemmin hoitaneiden ihmisten pitää etsiä uusia tehtäviä.

Tällaisten isojen sosiaalisten muutosten johdosta on tekoälysuunnittelussa pyrittävä aiempaa kokonaisvaltaisempaan ajatteluun. Ei riitä, että ihmisen käyttöön suunnitellaan älykäs tekninen laite. Olennaista on se, että toiminnan muuttumista katsotaan kokonaisuutena. Organisaatioita on kehitettävä, ja työntekijät johtoa myöten on koulutettava uusiin toimintamalleihin.

Tässä yhteydessä organisaatio ei tarkoita pelkästään yksittäistä yritystä tai julkishallinnon toimijaa, vaan asiaa on syytä miettiä myös kansakunnan tasolla. Uuden teknisen kulttuurin synty vaatii todella paljon innovaatiojärjestelmän eri osilta. Yritykset ja organisaatiot joutuvat arvioimaan uudelleen omat toimintamallinsa ja kehittämään niitä aikaa vastaavalla tavalla.

Toimintojen organisoiminen ei enää ole vain teknisen koulutuksen saaneiden ihmisten asia. Hyvä ohjelmoija ei juuri koskaan ole hyvä hallintotieteilijä tai organisaatiokehittäjä. Tämän vuoksi uudistuksen tekeminen edellyttää monialaisia tiimejä, joiden jäsenet ovat eri alojen asiantuntijoita.

Muutokset ovat välttämättömiä. Jos maamme ei kykene toimimaan luovalla tavalla, on varmaa, että esimerkiksi aasialaiset kilpailijat valtaavat markkinat. Suomalaisten ongelmana on yleensä kotimaisen skaalan pienuus. Tämän johdosta on tärkeää olla aivan kärjessä, sillä suurien resurssien avulla kilpailijat voivat viedä markkinat. Kärkiosaaminen suojaa kuitenkin toimintoja.

Suurin este järkeville muutoksille on ihmisten mielissä. On helppo juuttua vanhoihin ajattelumalleihin ja pyrkiä jarruttamaan kehitystä. Ihminen voi uskoa, että maailma säilyy ennallaan, vaikka tekoäly tunkeutuu pala palalta jokaiselle toimintasektorille. Muutokset tuntuvat monista epämukavilta, koska ne vaativat ponnistelua. Historiassa on kuitenkin melkoinen määrä esimerkkejä siitä, miten luovan ja innovatiivisen ajattelun pysähtyminen voi johtaa vain taantumiseen. Tietenkin monet uudistusyritykset osoittautuvat välillä harhalaukauksiksi. Tätä tosiseikkaa on helppo käyttää motiivina, jos ei halua mukautua välttämättömiin muutoksiin. Tätä tosiseikkaa on helppo käyttää motiivina pyrkimykselle olla mukautumatta. Maailma kuitenkin muuttuu, eikä aika jää odottamaan ketään.

8 Tekoäly kansalaisten palveluksessa

Tommi Mikkonen
professori
Jyväskylän yliopisto & Helsingin yliopisto

8.1 Johdanto

Tekoälyä voidaan ajatella tietokoneohjelmana, joka suorittaa älykkäinä pidettäviä toimintoja. Melko yleisesti hyväksytyn määritelmän mukaan tekoälyä on järjestelmän kyky tulkita ulkoisia tietoja oikein, oppia tällaisista tiedoista ja käyttää opittuja asioita tiettyjen tavoitteiden ja tehtävien saavuttamisessa joustavan sopeutumisen kautta[1]. Koska määritelmä jättää auki sen, missä mittakaavassa järjestelmä oppii asioita ja missä määrin se sopeutuu uusiin asioihin, tulkinnasta riippuen voimme nähdä tekoälyä oikeastaan kaikkialla tai sitten vain hyvin spesifeissä sovelluskohteissa. Toisaalta emme aina edes tiedä, mikä on tekoälyä, vaikka sellaista kohtaisimmekin.

Tässä artikkelissa tekoälyä ajatellaan laajasti, ja se on tarkoitettu kattamaan tekoälyn koko kirjo. Sen toisessa ääripäässä on automaattinen tietojenkäsittely, joka perustuu älykkääseen ohjelmaan. Sen toisessa päässä taas on tieteellisessä mielessä hyväksytty tekoäly eli hermoverkot ja koneoppiminen. Raportin tavoitteena on rakentaa tilannekuva siitä, millä tavalla tekoälyä voitaisiin hyödyntää julkisen sektorin palveluiden digitalisoinnissa kansalaisen näkökulmasta, sekä kuvata tämänhetkistä tietojärjestelmäkokonaisuutta julkisella sektorilla. Lisäksi raportissa esitetään kolme erilaista vaihtoehtoa, joiden avulla tekoäly voisi palvella kansalaisia aiempaa ketterämmin ja tasapuolisemmin.

Raportin tilannekuvaosion muodostavat lyhyet kuvaukset kansalaisten julkisiin palveluihin kohdistuvista tarpeista, julkisten palveluiden tekoälyllistämisestä sekä julkisen sektorin tietojärjestelmien tilasta Suomessa. Tämän jälkeen esitellään vaihtoehtoiset skenaariot, joihin Suomi voi turvautua tekoälyn hyödyntämisessä. Lopuksi raportti päättyy lyhyeen yhteenvetoon.

8.2 Tilannekuva

8.2.1 Kansalaisten tarpeet

Suomen kansalaiset – tai laajemmin ajateltuna myös Suomessa asuvat muiden maiden kansalaiset – muodostavat heterogeenisen ryhmän. Siksi on selvää, että he tarvitsevat ja toivovat hyvin erilaisia julkisen sektorin palveluita. Yhtäältä joukossa on tietotyön ammattilaisia, joille tietokoneiden ja muiden digitaalisten laitteiden käyttö on arkipäivää

1 Andreas Kaplan; Michael Haenlein (2019) *Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence*, Business Horizons, 62(1), 15–25.

ja vaivatonta, ja toisaalta osalla käyttäjistä ei vastaavia taitoja ole juuri lainkaan. Lisäksi on otettava huomioon erilaiset henkilökohtaiset ominaisuudet, kuten vaikkapa näkö- tai kuulo-ongelmat. Ne asettavat julkisen sektorin ICT-palveluille vaatimuksia, joita on vaikeaa toteuttaa täydellisesti. Myös palvelukokemus voi vaihdella sen mukaan, kokeeko palvelun käyttäjä saavansa palvelua ihmiseltä vai tietokoneelta.

Kaikkien mahdollisten kansalaisryhmien yhtäläinen huomiointi julkisen sektorin digitalisoinnilla ei ole mahdollista. Jos digipalveluita kuitenkin kohdennetaan sopivasti, voidaan vähentää merkittävästi kansalaisten henkilökohtaista palvelua. Silloin nämä resurssit voidaan ohjata sellaisiin palveluihin, joiden digitalisointi on vaikeaa. Pohjimmiltaan kyseessä on digitaalinen kuilu (engl. *digital divide*[2]). Se tarkoittaa sitä, että osa kansalaisista pystyy hoitamaan palvelutarpeensa nopeasti ja ketterästi itse, kun taas jotkut toiset ovat väistämättä ihmisien tarjoamien palveluiden varassa. Toisaalta jos palveluita kohdennetaan älykkäästi, yksittäinen kansalainen saa parempaa palvelua. Tämä johtuu siitä, että kansalainen voi hoitaa yhden (digitaalisen) palvelun kautta usean hallinnonalan asioita eikä hänen tarvitse hoitaa jokaista yksittäistä asiaa siihen erikoistuneen julkisen sektorin toimijan kanssa. Tällainen ihmis- tai kansalaiskeskeisyys (engl. *citizen-centricity*) edustaa uutta ajattelua, jota on kokeiltu myös julkisella sektorilla, esimerkiksi AuroraAI-ohjelmassa.[3]

Yksi ihmiskeskeisyyden ilmentymistä on AuroraAI:ssa käytetty tilannekuvamalli, jossa palveluita rakennetaan tietyn kohderyhmän johonkin tarkasti identifioituun tarpeeseen. Kun tilannekuva ja siihen liittyvät tarpeet tunnetaan, on mahdollista rakentaa tälle kohderyhmälle pilotteja juuri oikeaan tarpeeseen. Toinen tärkeä ihmiskeskeisyyden teema on se, että palveluiden on tarkoitus tukea kansalaisia yksilötasolla eikä pakottaa heitä johonkin tiettyyn muottiin.

8.2.2 Kansalaisille tarjotut tietojärjestelmät ja niiden data

Tekoäly kaikissa muodoissaan tarvitsee lähtökohdakseen dataa. Tämä data voi koostua oletuksista, simulaatioista tai todellisesta mittausdatasta. Jotta dataa voisi hyödyntää mahdollisimman vaivattomasti, sen pitäisi olla helposti käsiteltävässä, mielellään yhdenmukaisessa muodossa. Julkisissa palveluissa tämä data on tallennettuna erilaisiin tietojärjestelmiin.

Suomalaisella julkisella sektorilla on tyypillisesti ollut laaja paikallinen autonomia. Tästä syystä monista palveluista on toteutettu useita kilpailevia versioita, joiden sisältämä data ei ole suoraan keskenään vertailukelpoista tai yhteensopivaa. Lisäksi paikallisin päätöksin hankitut tietojärjestelmät saattavat johtaa hajanaisiin palveluihin: palvelut kattavat juuri hankkijan tarpeet, mutta esimerkiksi yhteensopivuusvaatimuksia tai muiden toimijoiden vastaavia tarpeita ei ole pohdittu. Esimerkkinä voidaan mainita Varsinais-Suomen hyvinvointialueen 41 (osin) päällekkäistä tietojärjestelmää.[4]

---

2 Wikipedia: Digital divide -sivu.
3 Valtiovarainministeriö (2020) Kansallinen tekoälyohjelma AuroraAI.
4 Viivi Laakkonen (2021): *Maakunnassa on 41 eri sote-alan tietojärjestelmää – kukaan ei tiedä, miten niistä luodaan yhtenäinen järjestelmä*. Turun Sanomat, 4.10.2021.

Tietojärjestelmiä on toteutettu useita erilaisia tekniikoita käyttäen. Osa palveluista on selvästi yleiskäyttöiseksi tarkoitettuja yksittäisiä palveluita, joiden varaan muut voivat rakentaa uusia palveluita. Hyvä esimerkki tällaisesta on Kansallinen terveysarkisto Kanta[5]. Osa järjestelmistä puolestaan on hankittu omina kokonaisuuksinaan, jotka tarjoavat palveluita loppukäyttäjille. Esimerkki tällaisesta on vaikkapa Apotti-järjestelmä[6]. Jälkimmäiset ovat usein suljettuja siinä merkityksessä, että niiden sisältämä tieto ei välttämättä ole käytettävissä muissa sovelluksissa. Setälä ja hänen kumppaninsa (2021)[7] esittävät näiden kahden lisäksi kolmannen mallin, jossa myös sovellukset rakennetaan uudelleen käytettävistä palvelukokonaisuuksista. Tämä malli näyttää kuitenkin olevan harvinainen käytännössä.

Valittu malli vaikuttaa mahdollisesti siihen, miten helppoa datapohjaisia ominaisuuksia on toteuttaa. Jos kaikki tieto on yhden järjestelmän sisällä, datan yhdenmukaisuus on helpompi taata. Toisaalta eri järjestelmien tietoja yhdistelemällä voidaan usein tehdä pidemmälle meneviä johtopäätöksiä kansalaisten tarpeista ja palveluista, joita heille pitäisi tarjota. Yhdistelyssä on kuitenkin huomioitava yksityisyydensuoja ja muut yksilöiden tietojen käsittelyyn liittyvät seikat, mukaan lukien EU:n yleinen tietosuoja-asetus GDPR.[8]

Joissakin palveluissa yksityinen sektori on ollut julkisen sektorin edellä. Tämä voi johtaa tilanteisiin, jotka eivät ole kansalaisten kannalta toivottavia. Esimerkki tällaisesta on vaikkapa pankkien tunnistautumisjärjestelmien käyttö kansalaisten digitaalisen identiteetin hallintaan – vaikka kansalaisten identiteetin hallinnan voisi helposti odottaa olevan valtion tasoinen asia. Monessa tilanteessa kaupalliset toimijat voivat perustellusti ottaa haltuunsa joitakin palveluita, mutta valtion tulisi pitää huolta kansalaisten kannalta strategisista tiedoista. Tällöin valtion ja mahdollisesti myös EU:n pitäisi kyetä koordinoimaan julkisen sektorin tietojärjestelmiä ja niihin liittyviä automaattisia ja tekoälypohjaisia päätöksentekoprosesseja ja sovelluksia, jotka hyödyntävät niitä. Parhaassa tapauksessa voisi syntyä oikea eurooppalainen markkina julkisen sektorin ohjelmistoille, samaan tapaan kuin mitä GAIA-X[9] tavoittelee teknisellä tasolla (GAIA-X, 2021).

8.3 Vaihtoehtoiset etenemistavat

8.3.1 Vaihtoehto 1: Paikalliset pilotit

Yksinkertaisin vaihtoehto tuoda tekoäly kansalaisten arkeen on antaa julkisille toimijoille vapaat kädet tehdä pilotteja kohdennetusti ja juuri oikeaan tarpeeseen. Kuten aiemmin todettiin, tällaisia pilotteja on jo toteutettukin AuroraAI-ohjelmassa, ja varmaankin melkein jokainen julkinen toimija on ainakin harkinnut tällaisia aktiviteetteja. Tässä

5 Wikipedia: Kansallinen Terveysarkisto -sivu.
6 Wikipedia: Apotti (potilastietojärjestelmä) -sivu.
7 Manu Setälä; Pekka Abrahamsson; Tommi Mikkonen (2021): *Elements of Sustainability for Public Sector Software – Mosaic Enterprise Architecture, Macroservices, and Low-Code*. In Proceedings of 12th International Conference on Software Business (2–3.12.2021, University of South-Eastern Norway, Norja). Springer.
8 Tietosuojavaltuutetun toimisto (2021) GDPR: EU:n tietosuoja-asetus.
9 GAIA-X:n kotisivu.

vaihtoehdossa julkinen toimija on keskiössä ja ratkaisut rakennetaan paikallisen tilanteen mukaan. Voidaan jopa ajatella, että paikalliset ratkaisut olisivat tavallaan selkeä jatko nykyiselle autonomialle. Toisaalta paikalliset pilotit johtavat helposti tilanteeseen, missä myös edut jäävät paikallisiksi. Silloin varsinkaan palveluiden tasapuolinen tarjoaminen kansalaisille ei toteudu.

Hyvät puolet:

- Palvelut kehitetään juuri oikeaan tarpeeseen.
- Pilotit on helppo muuttaa käytännöiksi paikallisesti.
- Palveluiden kehittäminen ei vaadi keskitettyä ponnistusta, vaan jokainen julkinen toimija päättää itse tarpeestaan ja toimistaan.

Huonot puolet:

- Opit eivät välttämättä leviä paikkakunnalta tai alueelta toiselle.
- Palvelut ovat tarjolla epätasa-arvoisesti.
- Samasta palvelusta voi syntyä päällekkäisiä toteutuksia.

8.3.2 Vaihtoehto 2: Teknologia- ja toteutusvetoiset toimenpiteet

Voidaan myös hyväksyä, että olemassa olevat järjestelmät ovat tekoälypohjaisten palveluiden lähtökohta. Sekin tarjoaa keinoja edetä kohti kansalaisille paremmin kohdennettuja palveluita. Tällöin tietolähteet – kuten aiemmin mainittu Kanta – ovat palveluiden rakentamisen ytimessä ja niiden ympärille rakennetaan uusia palveluita joko suoraan olemassa olevan tiedon varaan tai yhdistelemällä olemassa olevaa tietoa uudella, innovatiivisella tavalla. Tällä hetkellä tilanne on se, että palveluita hyödynnetään yksittäisissä hankkeissa ja kokeiluissa. Niitä ei kuitenkaan koordinoida mitenkään, eikä julkisilla toimijoilla ole yhteistä näkemystä siitä, miten edetä. Parhaassa tapauksessa tietojärjestelmät tarjoavat myös vientimahdollisuuden, jos palveluyritykset saadaan mukaan muodostamaan vientiverkosto ja tarjolla on riittävän suuri yhteinen kokonaisuus.

Hyvät puolet:

- Palvelut perustuvat olemassa oleviin järjestelmiin, joten tarvittavat investoinnit ovat pieniä.
- Jokaisella tietojärjestelmällä on omistaja, joka voidaan vastuuttaa tarjoamaan tiedot tai vähintään rajapinta muille.
- Palveluiden kehittäminen ei vaadi keskitettyä ponnistusta, vaan jokainen julkinen toimija päättää itse tarpeestaan ja toimistaan.

Huonot puolet:

- Tietojärjestelmien hyödyntämisessä ei voida käyttää yhtä ainoaa reseptiä, vaan tilannetta pitää tarkastella tietojärjestelmäkohtaisesti sekä teknisessä että lainsäädännöllisessä mielessä.
- Tiedon luottamuksellisuus voi estää sen yhdistämisen muihin tietolähteisiin tai jopa tehdä tiedosta puhtaasti yhdelle toimijalle kuuluvaa.
- Kansallinen ohjaus uusien toteuttamisessa jää puutteelliseksi.

8.3.3 Vaihtoehto 3: Strategiset toimenpiteet

Kolmas vaihtoehto on siirtää julkisten tietojärjestelmien toteutus valtion strategiseen ohjaukseen. Valtion tasolla voidaan luoda kansallinen viitearkkitehtuuri ja tiekartta siitä, missä järjestyksessä eri palvelut integroidaan viitearkkitehtuuriin. Samalla voidaan pohtia, miten edetä pilotoinnissa. Tavallaan tämä vaihtoehto yhdistää kaksi edellistä, mutta niin, että niillä on keskitetty johto. Tämä johto puolestaan voi ottaa huomioon EU:n digitaalisen kompassin ja ohjata julkista sektoria kohti niitä tekoälyn hyödyntämistä koskevia tavoitteita, jotka on määritelty EU:n tasolla.

Hyvät puolet:

- Valtion ohjaus mahdollistaa ohjelmistojen uudelleen- ja yhteiskäytön. Esimerkiksi tulevat hyvinvointialueet voisivat hyödyntää samaa ohjelmistoa, eikä jokaisen tarvitsisi rakentaa omaansa.
- Kun kehitystyötä koordinoidaan, päätöksentekoa ja tiedonjakoa voidaan keskittää
- Palveluiden saatavuus on tasapuolisempaa kuin silloin, jos jokainen julkinen toimija saa itse päättää toimistaan.
- Pitkällä aikavälillä ohjelmistojen uudelleenkäyttö laskee kustannuksia.

Huonot puolet:

- Keskeiset valtion tason toimijat ovat ylityöllistettyjä.
- Tietojärjestelmien strateginen johtaminen on vaikeaa, sillä siihen soveltuvia mittareita on niukasti tarjolla.
- Paikallisten toimijoiden autonomia on vaikea purkaa.
- Keskittäminen lisää lyhyellä tähtäimellä kustannuksia.
- Valtion ohjaus vaatii kulttuurin ja varmaankin myös organisaation muutosta, sillä se ylittää perinteiset organisaatiorajat.

8.4 Yhteenveto ja suosituksia

Suomen julkisen sektorin tietojärjestelmien kokonaistilaa on vaikea arvioida, sillä toimijoita on useita suuren paikallisen autonomian vuoksi. Tästä syystä myös tekoälyn mahdollisuudet ovat erilaiset, mitä voidaan pitää jossain mielessä epätasa-arvoisena. Toisaalta palveluiden paikallinen kohdentaminen mahdollistaa juuri oikeaan tarpeeseen toteutetut palvelut, koska silloin voidaan ottaa huomioon myös paikalliset edellytykset. Kokonaisuutena ajatellen julkisen sektorin eri toimijat kuitenkin tekevät helposti päällekkäistä työtä ja hankkivat samankaltaisia palveluita yhä uudelleen ja uudelleen. Pahimmassa tapauksessa nämä hankinnat eivät edes hyödynnä jo tehtyä työtä, vaan uusi järjestelmä hankitaan kokonaan uutena toteutuksena, joka rakennetaan juuri tilaavaa toimijaa varten, juuri tilaajan tarpeeseen. Tällöin käy helposti niin, että iso kuva tietojärjestelmistä ja niiden yhteensopivuudesta jää huomioimatta.

Jos tilannetta halutaan muuttaa, tarvitaan valtion tason toimenpiteitä, joilla parannetaan järjestelmien itsensä sekä niiden sisältämän datan yhteensopivuutta. Tämän prosessin keskiössä on pyrkimys uudelleen käytettäviin palveluihin ja niiden päälle rakennettaviin sovelluksiin, sen sijaan että ostetaan kokonaisia järjestelmiä avaimet käteen -periaatteella. Jotta tässä voidaan onnistua, tarvitaan strategista visiota siirtymässä kohti kansalaisia

palvelevaa tekoälyä sekä johtajuutta ja rohkeutta lähteä toteuttamaan visiota yli perinteisten organisaatiorajojen.

Kansalaisten kannalta tämä strateginen johtajuus mahdollistaa sen, että tekoälypohjaiset palvelut ovat tasa-arvoisia ja että ne on toteutettu tavalla, joka ottaa huomioon yksityisyydensuojan. Lisäksi palvelut pitää toteuttaa siten, että ne tukevat ihmisiä – ei niin, että ne pakottavat yksilöt toimimaan järjestelmän ehdoilla.

Lisäksi palveluita kannattaisi ohjata koko Euroopan tasolla, jotta voidaan luoda yhtenäinen markkina julkiselle sektorille. Tämä on myös Suomen etu, sillä suomalaiset ohjelmistoyritykset ovat toteuttaneet useita laadukkaita ohjelmistoja julkiselle sektorille. Ne voidaan muuttaa vientituotteiksi, jos palvelut ovat yhdenmukaisia Euroopan maissa.

Suosituksia:

- Suomi tukee EU:n digitaalista kompassia ja ottaa asian huomioon valtion tason IT-toiminnassa nimeämällä digitalisaatioministerin tai vahvistamalla ICT-johtajan roolia.
- ICT:n kehityksessä asetetaan tavoitteeksi kansallinen strateginen ohjaus, jossa muodostetaan yhteinen kansallinen viitearkkitehtuuri ja tiekartta siitä, missä järjestyksessä eri palvelut integroidaan viitearkkitehtuuriin.
- Selvitetään mahdollisuudet julkisen sektorin ohjelmistoviennille ja sitä kautta suomalaisen tekoälyn hyödyntämiselle laajasti Euroopassa.
- Tekoälyä koskeva lainsäädäntö pitää laatia siten, että se on teknologiasta riippumatonta ja koskee kaikkea tietojenkäsittelyä. Muuten jää paljon harmaata aluetta sen suhteen, mikä on tekoälyä ja mikä ei.

9 Tekoäly terveydenhuollossa

Mark van Gils
professori
Tampereen yliopisto

9.1 Mistä on kyse?

Terveydenhuolto on tekoälyteknologioiden luonnollinen sovellusalue. Terveydenhuollossa on monia vaativia tai aikaa vieviä tehtäviä, prosesseja ja toimintoja, joissa kaikki apu on enemmän kuin tervetullutta. Kun hoitojärjestelmä ja -henkilökunta ovat kasvavan paineen alla, apua haetaan yhä enemmän teknologisista ratkaisuista. Niiden joukossa tekoäly on avainasemassa.

Tekoälystä voi olla hyötyä monenlaisissa terveydenhuoltopalveluissa. Alla on muutamia esimerkkejä tilanteista, joissa tekoälyä voi hyödyntää:

- Neurologin on tehtävä diagnoosi potilaasta, jolla on epäselviä, muistinmenetykseen viittaavia oireita.
- Lääkärin on päätettävä, onko turvallista tehdä leikkaus henkilölle, joka saapui juuri ensiapuun auto-onnettomuuden vuoksi.
- Perusterveydenhuollon lääkärin on päätettävä, lähetetäänkö potilas erikoislääkärille.
- Radiologin on tulkittava lääketieteellinen kuva.
- Sairaanhoitaja seuraa potilaan tilaa osastolla.
- Keski-ikäinen terve mies haluaa pysyä terveenä.
- Henkilö, jolla on todettu tyypin 2 diabeteksen korkea riski, haluaa vähentää riskitekijöitä.
- Vanhuksen aikuiset lapset haluavat seurata etänä, onko vanhemmalla kaikki kunnossa.
- Sairaalan johtoryhmä haluaa seurata osastojen tehokkuutta.
- Osastonhallinnon on suunniteltava resurssien käyttöä.
- Päättäjät arvioivat suunnitellun lainsäädännön tai investointien mahdollisia vaikutuksia.

Monet näistä tapauksista ovat monimutkaisia. Niiden selvittämiseksi käytetään useita erilaisia tietolähteitä, mutta aina niihin ei ole vain yhtä oikeaa ratkaisua. Niillä kaikilla on merkittäviä seurauksia esimerkiksi ihmisten terveydelle, yhteiskunnalle aiheutuville kustannuksille sekä potilaiden, omaisten ja terveydenhuollon henkilöstön hyvinvoinnille. Ajatus tekoälyn käyttämisestä terveydenhuollon päätöksenteon tukena ei ole uusi.

Ohjeita ja lääketieteellistä asiantuntemusta on käytetty sääntöpohjaisissa asiantuntijajärjestelmissä 1970-luvulta lähtien. Melko pian kuitenkin huomattiin, että nämä järjestelmät eivät kata kaikkea sitä hiljaista tietoa, jota asiantuntijat käyttävät työssään. Niitä ei

voida myöskään päivittää helposti, kun alalla muodostuu uutta tietoa. Koska datan saatavuus, laskentateho ja datapohjaiset tekoälyalgoritmit ovat kehittyneet terveydenhuolto-alalla merkittävästi, ”tekoälystä terveydenhuollossa” on tullut synonyymi datapohjaiselle tekoälylle. On kuitenkin hyvä muistaa, että säännöt ja ohjeet ovat edelleen tarpeen esimerkiksi hoitosuositusten ja käytännön ohjeiden kohdalla.

Tekemässään selvityksessä Korhonen, Ermes ja Ahola[1] antoivat yleiskuvan tekoälyn mahdollisuuksista terveydenhuollossa Suomessa. He tunnistivat kuusi aluetta, joilla on odotettavissa eniten hyötyjä. Nämä ovat ensinnäkin 1) personoitu hoito, 2) automaattinen terveystietoanalyysi ja 3) jatkuva kansalaiskeskeinen hoito. Näitä kolmea aluetta ohjaavat erityisesti Suomessa olevat tietovarannot, resurssit ja osaaminen, ja niitä tukevat teollisuuden vahvat intressit. Kolme muuta osa-aluetta ovat 4) terveydenhuoltoprosessien- ja sosiaalipalveluprosessien kehittäminen, 5) palvelujen automatisointi terveydenhuollossa ja sosiaalihuollossa ja 6) tietoon perustuvat kansanterveystyön päätösmekanismit. Näillä alueilla tekoälyn hyödyntäminen parantaa terveydenhoidon ja sosiaalitoimen tuottavuutta ja siten pienentää kustannuspaineita.

Sittemmin automaattinen terveystietoanalyysi on ottanut selkeitä edistysaskeleita varsinkin tehtävissä, joita voidaan pitää ”kapeana tekoälynä”. Hyvänä esimerkkinä tästä on lääketieteellisten kuvien automaattinen käsittely ja tulkinta lääketieteellisissä laitteissa tai järjestelmissä. Jatkuva kansalaiskeskeinen hoito, joka sisältää esimerkiksi henkilön fysiologisen tai vireystilan päivittäisen seurannan, on kasvanut nopeasti. Konkreettisesti tämä näkyy esimerkiksi älykellojen ja -sormusten suosiossa. Tutkimus laajenee voimakkaasti personoidun hoidon ja lääketieteen aloilla muun muassa biopankkien ja yleisöltä kerätyn datan ansiosta.

Toisessa ryhmässä (osa-alueet 4–6) konkreettisimmat askeleet on otettu palveluautomaatiossa, johon ”kapea tekoäly” sopii parhaiten. Esimerkiksi chatbotit hoitavat ensikontaktin potilaiden ja terveydenhoidon välillä. Chatbot-automaatio perustuu puheentunnistuksessa ja luonnollisen kielen käsittelyssä (NLP) saavutettuun edistykseen. Kaksi muuta aluetta eli prosessien kehittäminen ja tietoon perustuva päätöksenteko julkista terveydenhoitoa ja kansanterveyttä koskevissa asioissa ovat luonteeltaan huomattavasti monimutkaisempia. Kuten vuonna 2017 ennustettiin, ne tuottavat tuloksia vasta pidemmällä aikavälillä. Tämä ei ole tilanne pelkästään Suomessa. Esimerkiksi äskettäin julkaistu tietokanta *FDA:n hyväksymistä tekoälyä hyödyntävistä* lääkinnällisistä laitteista osoittaa, että suurin osuus koskee radiologian, kardiologian, endokrinologian ja onkologian laitteiden diagnostiikkatehtäviä. On myös syytä todeta, että covid-19-pandemian myötä etälääketieteen hyödyntäminen sai melkoisen sysäyksen. Etälääketiede, kuten telekonsultointi ja etävalvonta, otettiin rutiinikäyttöön tehtävissä, joissa se aiemmin nähtiin enimmäkseen tulevaisuuden lupauksena.

1 Korhonen, Ilkka, Miikka Ermes ja Jari Ahola. 2017. **Strategic Research Agenda (SRA) on “*Finnish Innovation Hub for Artificial Intelligence for Health (AI for Health)*.”** VTT Technology. Espoo: VTT Technical Research Centre of Finland.

Tekoälyn käyttöönotto terveydenhuollossa on kuitenkin hidasta verrattuna esimerkiksi teollisuuteen, liikenteeseen, finanssialaan ja mediasektoriin. Haasteeksi muodostuu ainakin kolme asiaa:

1. Terveydenhuollon data on erityislaatuista. Se on usein kohinaista, osin virheellistä ja yhteensopimatonta. Tämän vuoksi on vaikeaa hankkia riittävän laadukasta dataa tekoälyn opettamista varten.
2. Tekoälyalgoritmien opettamiseen liittyy terveydenhoidossa omat vaikeutensa, koska usein ei ole olemassa selvää ja yksinkertaista ”oikeaa tulosta”, jota tarvittaisiin tekoälyn opettamisessa. Mittausdatan taustalla oleva tauti tai tieto potilaan tilan paranemisesta tai huononemisesta saadaan selville usein vasta monta vuotta mittauksia myöhemmin. Tämä voi johtua siitä, että dataan liittyviä selventäviä merkintöjä on niukasti, potilaan diagnoosi ei ole absoluuttisen oikea tai terveydenhuollon asiantuntijat ovat erimielisiä potilaan tilasta tai sen syistä.
3. Päätöstukijärjestelmän ottaminen jokapäiväiseen käyttöön edellyttää sopeutumista lukuisiin ei-teknisiin vaatimuksiin ja oletuksiin sekä niiden noudattamista. Vakiintuneet prosessit, asenteet ja toimintakulttuuri vaativat, että järjestelmä sopeutuu niihin.

Lainsäädännölliset, eettiset ja yksityisyydensuojaan liittyvät näkökohdat ja sääntelykysymykset vaikuttavat kaikkeen. Näiden haasteiden ratkaisemiseksi tarvittavat ponnistelut aliarvioidaan usein karkeasti, koska ne edellyttävät todella monialaista ymmärrystä ja eri tahojen sitoutumista.

Tekoälyyn liittyvä tutkimus voidaan karkeasti jaotella kolmeen ryhmään, jotka vastaavat edellä kuvattuihin kolmeen haasteeseen. Tutkimuksella haetaan vastausta kysymykseen, kuinka käsitellä tehokkaammin huonolaatuista tai puutteellista dataa ja jäsentämättömiä tietoja. Jälkimmäiseen liittyviä menetelmiä ovat ohjaamaton ja puoliohjattu opetus (engl. *unsupervised and semi-supervised training)*. Kuva-analyysin kehittäminen jatkuu voimakkaasti ja laajenee lääketieteellisen kuvantamisen sovelluksista uusiin alueisiin. Eri tietolähteiden yhdistäminen on välttämätöntä, jotta saadaan kokonaisvaltainen näkemys potilaasta, ja tämä antaa erityisen panoksen tekoälyn lähestymistapoihin, joissa käytetään monenlaisia datalähteitä, kuvista ja –”omiikka datoista” strukturoimattomiin muistiinpanoihin. Merkittäviä tutkimuspanoksia kohdistuu datassa tai algoritmeissa olevien vääristymien (engl. *bias*) havaitsemiseen ja poistamiseen, yksityisyyden säilyttämiseen sekä tekoälymenetelmien luotettavuuteen, läpinäkyvyyteen ja eettisiin kysymyksiin. Ponnistelua tarvitaan myös terveyslukutaidon parantamiseksi ja terveyteen liittyvän disinformaation torjumiseksi – esimerkkinä rokotteisiin liittyvä disinformaatio. Tekoäly voi auttaa myös näissä asioissa. Vaikka monet näistä aiheista eivät sinänsä ole ominaisia vain terveydenhuollon alalle, niillä on suurempi rooli terveydenhuollossa kuin monilla muilla sovellusalueilla.

Näiden alueiden lisäksi vahvaa tutkimusta tehdään muun muassa lääkekehityksessä, lisätyssä ja virtuaalitodellisuudessa (esim. koulutus, kuntoutus ja kirurgia), eksoskeletoneissa sekä henkilökuntaa avustavassa ja potilaiden kuntoutusta tukevassa robotiikassa.

9.2 Tekoäly muutosvoimana

Lyhyellä ja keskipitkällä aikavälillä tekoälyä hyödynnetään terveydenhoidon osa-alueilla, joissa on jo päästy liikkeelle. Näitä ovat a) diagnostiikka ja hoidon suunnittelu varsinkin kohteissa, joissa dataa on runsaasti saatavilla, ja b) työläiden, rutiininomaisten tehtävien keventäminen automaatiolla. Näitä ovat ”back-office-automatisointi” toiminnan tehostamiseksi: automaattinen resurssisuunnittelu, puheen muuttaminen tekstiksi ja muistiinpanojen tuottaminen sekä logististen prosessien optimointi. Robotiikkaa on kehitetty myös logistiikkaa ja prosessien optimointia varten. Lääketeollisuudessa puolestaan hyödynnetään lohkoketjuteknologioita toimitusketjun hallinnassa. Potilaan hoitopolkujen tukena käytetään chatbotteja. Konsulttiyritys McKinsey näkee *kansainvälisillä markkinoilla* tekoälyn kannalta lupaaviksi seuraavat osa-alueet: tarkkuuslääketiede, etäpotilastuki ja digitaaliset apteekit ja päätöksenteko. Kuten edellä mainittiin, etälääketiede on nyt tulossa rutiinikäyttöön. Ennen vuotta 2020 telelääketieteen käyntien osuus oli maailmanlaajuisesti alle kaksi prosenttia kaikista avohoidon käynneistä. Osuus on noussut 30 prosenttiin covid-19-pandemian aikana. Tekoälypohjaiset älykkäät kotivalvontajärjestelmät, etäkonsultointi ja etävalvonta ovat lisääntyneet voimakkaasti. Nyt, kun nämä tekniikat ovat tulleet tutuiksi ja hyväksytyiksi, niiden käyttö laajenee yhä.

Väestörakenteen muutos, erityisesti odotetun eliniän piteneminen, lisää kalliiden kroonisten sairauksien esiintyvyyttä. Niiden ehkäiseminen tai ainakin viivästyttäminen on tehokkain tapa hillitä kustannuksia terveydenhuoltojärjestelmässä. Tämä korostaa erityisesti ennalta ehkäisevää hoitoa ja riskien vähentämistä muun muassa elämäntapamuutoksilla. Tekoälyä voidaan hyödyntää tukemaan elämäntapamuutoksia (ohjelmat ja sovellukset, älykellot, puettavat laitteet), etävalvontaa sekä etälääketieteen ratkaisuja.

Eri ikäryhmät käyttävät tekniikkaa eri tavoin. Myös monikulttuurisuus, yksilöllistyminen ja sosiaalinen syrjäytyminen ovat yhä näkyvämpiä ilmiöitä, jotka voivat monimutkaistaa tekoälypohjaisten teknisten ratkaisujen omaksumista ja soveltamista. Ne myös vaativat huolellista tarkastelua tekoälyn kehitys- ja käyttöönottovaiheessa. On siis tärkeää tutkia tekoälypohjaisia ratkaisuja vuorovaikutuksessa käyttäjien kanssa ja ottaa alusta asti huomioon erilaiset tarpeet. Ei ole yhtä ratkaisua, joka sopii kaikille.

Myös kestävän kehityksen näkökulmat on syytä ottaa huomioon. Tehokkaammat tekoälyalgoritmit, jotka tarvitsevat yhä tehokkaampaa laskentaa ja suurempia datamääriä, lisäävät suorituskykyä, mutta samalla niiden sähkönkäyttö ja hiilijalanjälki ovat eksponentiaalisesti suurempia[2]. Tämäntyyppiset näkökohdat vaikuttavat kustannustehokkuuden arviointiin. Meidän on pohdittava, millä kustannustehokkuudella ja ympäristövaikutuksilla saadaan esimerkiksi viiden prosentin lisätarkkuus diagnostiikkaan tekoälyn avulla, kun siihen tarvitaan valtava määrä dataa ja laskentaresursseja. Tarve katkaista suorituskyvyn riippuvuus hiilijalanjäljestä on yhä suurempi tekoälyä kehitettäessä.

---

2 Deep Learning’s Diminishing Returns. 2021. IEEE Spectrum. September 24, 2021.

9.3 Viestit ja suositukset päättäjille ja kansalaisille

Yhteenvetona voidaan todeta, että terveydenhuolto ja tekoäly muodostavat yhdistelmän, jolla on valtava potentiaali. Mutta tämän potentiaalin hyödyntäminen on monimutkaista ja vaativaa. Kehityshankkeiden tavoitteiden asettaminen ja tulosten arviointi ovat monimutkaisia, ja ympäristö, jossa tekoälymenetelmiä käytetään, on usein vieläkin kompleksisempi. Tästä voidaankin vetää johtopäätös, että tekoäly ei tuota pikavoittoja terveydenhuoltohankkeissa. Investointien on oltava pitkäjänteisiä, ja on hyväksyttävä, että ne alkavat tuottaa vasta monien vuosien kuluttua. Menestys edellyttää, että investointeja, rahoitusjärjestelyjä ja poliittista päätöksentekoa suunnitellaan pitkällä aikavälillä, koska panostusten hyödyt eivät ole välittömästi näkyvissä. Toinen edellytys onnistumisille on yhteistyö. Tutkijat, potilaat, ammattilaiset, teollisuus, poliittiset päättäjät ja sääntelyelimet on kaikki otettava huomioon alusta alkaen, jotta voidaan toivoa todellista menestystä.

Yhteisten aloitteiden tarve on tunnistettu hyvin, ja toimia toteutetaan. EU:n tutkimusohjelmilla (Horizon Europe), Business Finlandilla ja Suomen Akatemialla on tärkeitä välineitä, jotka on erityisesti tarkoitettu laajaan yhteistyöhön ja ekosysteemien kehittämiseen. Tietojen jakamisaloitteet ovat keskeisessä roolissa, ja ne ulottuvat kansallisista biopankkiyhteistöistä ja data-altaista (engl. *data lake*) laaja-alaisempiin toimiin, kuten eurooppalaiseen terveystietoa koskevaan aloitteeseen (European Health Data Space). Näihin liittyviä hankkeita ovat muun muassa Sitran *TEHDAS Joint Action* ja *GAIA-X, turvallinen datainfrahanke*, joka kattaa monia sektoreita, mukaan lukien terveyssektorin.

Kansainvälisen tason lainsäädäntö, kuten EU:n yleinen tietosuoja-asetus (GDPR), ja kansallinen terveystietojen toissijaista käyttöä koskeva laki avaavat mahdollisuuksia, vaikka niiden täytäntöönpanossa ja vaikutuksissa esiintyvien vielä horjuvien tulkintojen takia on haasteita, joihin tarvitaan resurssit ja lisää yhteistyötä. EU:n tekoälyä koskeva koordinaatiosuunnitelma vaikuttaa myös siihen, miten tekoälyä voidaan hyödyntää terveydenhuollossa. Tutkimusyhteisö suhtautuu koordinaatiosuunnitelmaan yleisesti myönteisesti, mutta siihen toivotaan silti tiettyjä parannuksia. Huolenaiheisiin lukeutuvat komission käyttämä hieman epämääräinen tekoälyn määritelmä sekä tutkimus- ja kehitystyöhön varatun rahoituksen määrä suhteessa tavoitteisiin. Lisäksi innovaatiojärjestelmiä on koordinoitava paremmin, ja ne on linkitettävä muiden tieteenalojen kokemuksiin (esim. ”Privacy and Safety by Design” -mallia voitaisiin käyttää esikuvana, kun luodaan tekoälyn eettisiä ohjeistuksia). Euroopan tekoälytutkimuksen laboratorioiden yhteistyöelimen (CLAIRE) jäsenet ovat antaneet *yksityiskohtaista palautetta koordinaatiosuunnitelmasta*, joka olisi hyvä huomioida.

Tärkeintä on kuitenkin se, mitä uusi tekoälypohjainen lähestymistapa saa terveydenhuollossa todella aikaan, olipa kyse sitten potilaiden hoitojaksojen pituudesta, euroista potilasta kohden vuodessa, potilastyytyväisyydestä, osastobudjetin käytöstä tai potilaan elämänlaadusta. Ratkaisevaa on, että vaikutukset voidaan mitata tehokkaasti ja määrittää numeerisesti. Mittaustuloksia on myös hyödynnettävä tavoiteprosessien ja -palvelujen parantamiseksi, riippumatta siitä, ovatko nämä prosessit ja palvelut tekoälypohjaisia vai eivät. Tämä on laajalti tunnustettu yhteisössä, mistä esimerkkinä ovat yliopistojen tutki-

mushankkeet[3] ja laajalti tuetut *suositukset* kansallisen sosiaali- ja terveydenhuollon vaikuttavuuskeskuksen perustamisesta[4].

Yhdistämällä nämä toimet meillä on kaikki mahdollisuudet hyödyntää tekoälyn valtava potentiaali terveydenhuollossa.

9.4 Terveydenhuolto ja tekoäly – kolme haastetta, joihin tarvitaan apua

Terveydenhuollon **datan erityislaatuisuus**: data on monesti kohinaista, osin virheellistä ja yhteensopimatonta. Se on usein myös vaikeasti saatavilla.

Usein **ei ole olemassa yksinkertaista "oikeaa tulosta"**, jota tarvittaisiin tekoälyn opettamisessa. Tulokset (esim. lopullinen diagnoosi, kustannukset tai potilaan elämän-laatu) voivat näkyä pitkällä viiveellä ja ovat hankalasti mitattavissa.

Päätöstukijärjestelmän o**ttaminen jokapäiväiseen käyttöön edellyttää sopeutumista lukuisiin ei-teknisiin oletuksiin ja vaatimuksiin sekä niiden noudattamista**. Organisaatioiden vakiintuneet prosessit, asenteet ja toimintakulttuurit vaikuttavat siihen, kuinka tekoälyn hyödyntämiseen suhtaudutaan. Lisäksi lainsäädännölliset sekä eetti-set ja yksityisyydensuojaan liittyvät näkökohdat ja sääntelykysymykset vaikuttavat kaikkeen.

3 Esim. Tampereen yliopiston tutkimusalusta *TRANSFORM (Towards Effective and Equitable Welfare Services: Translating Scientific Knowledge into Improved Practices and Outcomes)*

4 *Vaikuttavuuden vuosikymmentä rakentamassa – Kansallinen sosiaali- ja terveydenhuollon vaikuttavuuskeskus -selvitys*, Itä-Suomen yliopisto ja Helsingin yliopisto, 2021

10 Tekoäly kaupunkikehityksessä – ihmiskeskeisen suunnittelun näkökulma

Kaisa Väänänen
professori
Tampereen yliopisto

Kaupungistuminen on kiihtyvä megatrendi sekä maailmanlaajuisesti että Suomessa. Yli puolet maailman väestöstä asuu tällä hetkellä kaupungeissa, ja tämän osuuden ennustetaan kasvavan 60 prosenttiin vuoteen 2030 mennessä.[1] Tämä kehitys on vielä nopeampaa länsimaissa; noin 70 prosenttia Suomen väestöstä asuu tällä hetkellä kaupungeissa[2], ja vuoteen 2030 mennessä kaupunkilaisia on arviolta 87 prosenttia. Kiihtyvä muuttoliike kaupunkeihin luo kasvavia paineita kaupunkialueiden kestävälle kehittämiselle.

Samaan aikaan maailmassa on meneillään laajamittainen yhteiskunnan toimintojen digitalisaatio, joka muuttaa yritysten, hallintojen ja yksityisten ihmisten toimintatapoja. Digitalisaation osana käytetään yhä useammin tekoälyn tekniikoita. Tekoälyn hyödyntämisellä voidaan saavuttaa suuria etuja, joita ovat esimerkiksi säästöt resurssien käytön optimoinnista, päätöksenteon sujuvoituminen ja turvallisuuden lisääntyminen. Toisaalta tekoälyn käyttöön liittyy myös riskejä, kuten yksityisyyden loukkaukset, päätöksenteon vinoumat sekä työttömyyden ainakin väliaikainen lisääntyminen automaatiotason noustessa.

Tekoälyä on jo mukana monissa kaupunkien sovelluksissa, usein ilman, että kansalaiset ovat tietoisia ”älykkyyden” luonteesta. Tekoälyn hyödyntäminen pohjautuu laajoihin datamääriin, joita kerätään ihmisten toiminnasta kaupunkikontekstissa. Tekoäly voi vaikuttaa konkreettisesti kansalaisten elämään, joten tekoälyyn voi liittyä epävarmuuden tunnetta, huolenaiheita tai jopa pelkoja.[3] Jotta tulevaisuuden älykkäiden kaupunkien toimintaa ja kestävyyttä voitaisiin parantaa tekoälyä hyödyntävillä toiminnoilla, ihmiskeskeisen lähestymistavan tekoälyyn pitäisi ohjata tekoälysovellusten kehittämisprosessia.

1 YK, United Nations, Department of Economic and Social Affairs, Population Division (2018). *World Urbanisation Prospects: The 2018 Revision, Online Edition*. (27.10.2021)
2 Suomen ympäristökeskus Syke (2021). *Tutkimusta ja kokeiluja kestävän kaupungistumisen edistämiseksi.* (27.10.2021)
3 Lehtiö, A., Hartikainen, M., Ala-Luopa, S., Olsson, T., & Väänänen, K. (2021). Understanding Citizens Perceptions of AI in the Smart City. Journal of AI & Society: Special issue on Urban AI: Understanding the Emerging Role of Artificial Intelligence in Smart Cities, Springer. In press.

10.1 Mistä ihmiskeskeisessä tekoälysuunnittelussa on kyse?

Tekoälyä on tutkittu ja kehitetty isoin harppauksin jo 1950-luvulta lähtien. Viime vuosina on koettu tekoälyn niin sanottu kolmas aalto, jonka aikana yhä useampia tekoälyinnovaatioita on saatu vietyä laajalti käytäntöön ja hyödynnettyä eri sovelluksissa. Sovellukset mukautuvat käyttäjien toimintaan käyttäen isoja tietomääriä. Näitä tietoja kerätään erilaisilla laitteilla käyttöympäristöistä ja ihmisiltä. On kuitenkin merkittävää, että tekoälyn tutkimuksessa ja kehityksessä on ollut vahvasti tekninen painotus, ja esimerkiksi Suomi on yksi tekoälyalgoritmien kehityksen ja tutkimuksen johtavista maista. Kuten minkä tahansa ihmisen kanssa vuorovaikutuksessa olevan teknologian kehityksessä myös tekoälyn kohdalla tulisi kiinnittää huomiota sosioteknisiin kysymyksiin. Tekoälyllä vahvistettu teknologia on usein proaktiivinen toimija, joka voi käyttää valtaa päätöksenteossa ja toimintojen ohjaamisessa. Tekoälyn yhteydessä ihmisiin ja yhteiskuntaan liittyvät kysymykset ovat siksi jopa merkittävämpiä kuin niin sanottujen ei-älykkäiden teknologioiden kohdalla.

Ihmiskeskeinen tekoäly (engl. *Human-Centered AI)* tarkoittaa näkökulmaa, jossa tekoälyavusteisten järjestelmien kehityksen lähtökohdaksi otetaan ihmisten ja yhteiskunnan tarpeet.[4] Ihmiskeskeinen tekoäly tarjoaa teknisten järjestelmien suunnitteluun tulokulman, jossa keskitytään ensisijaisesti kansalaisten tarpeisiin, näkemyksiin, tekoälyn etiikkaan, teknologiapohjaisten palvelujen hyväksyntään ja laajempaan sosiokulttuuriseen kontekstiin. Näin tekoälyä lähestyttäessä voidaan tuottaa yhteiskuntaan toimivia ja hyväksyttäviä tekoälyjärjestelmiä.

Ihmiskeskeisessä tekoälysuunnittelussa on seuraavia erityispiirteitä:

- Ihmisten tarpeet otetaan kattavasti huomioon koko kehityshankkeen ajan. Järjestelmä- ja sovellussuunnittelun alkuvaiheesta alkaen selvitetään sekä järjestelmän loppukäyttäjien että muiden järjestelmän vaikutuspiirissä olevien asianosaisten (engl. *stakeholders*) tarpeet ja huolet. Nämä sosiotekniset ja psykologiset tarpeet muodostavat pohjan järjestelmän suunnittelulle yhdessä teknisten ratkaisujen kanssa. Järjestelmää pitää arvioida vaatimuksia vasten iteratiivisesti kehityksen aikana. Lisäksi pitää kartoittaa eettiset kysymykset ja niiden ratkaisut.
- Ihmisen ja tekoälyn vuorovaikutus suunnitellaan siten, että käyttäjä voi hallita tilannetta. Useissa tekoälyjärjestelmissä käyttäjä ja tekoälyjärjestelmä toimivat tietyllä tavalla yhteistyössä tai ovat vuoropuhelussa, jossa tekoäly tekee suosituksia tai alustavia päätöksiä. Käyttäjällä on kuitenkin oltava mahdollisuus osallistua päätöksentekoon. Toisin sanoen ihmisen ja koneen välisen vuorovaikutuksen avulla ihmisellä säilyy kontrolli tärkeissä toiminnoissa. Datan keruu ihmisiltä täytyy tuoda selkeästi esille, myös silloin, kun se tapahtuu julkisessa tilassa, kuten kaupungissa.

---

4 Schmidt A., Giannotti F., Mackay W., Shneiderman B., Väänänen K. (2021). *Artificial Intelligence for Humankind: A Panel on How to Create Truly Interactive and Human-Centered AI for the Benefit of Individuals and Society*. Proceedings of INTERACT 2021. LNCS, vol 12936. Springer, Cham.

- Tekoälyn ”mustan laatikon” kaltaista luonnetta valaistaan. Tekoälyn toiminta ei ole ihmisille helposti ymmärrettävää, joten ihmisen ja tekoälyn vuorovaikutuksessa on pyrittävä tekoälyn selitettävyyteen. Tekoälysovellusten käyttöliittymät pitää suunnitella siten, että käyttäjä saa riittävästi tietoa siitä, miten tekoäly on tehnyt suosituksensa tai päätöksensä.
- Sovelletaan ihmiskeskeisiä menetelmiä tekoälyjärjestelmien suunnitteluun. Jotta edellä kuvatut piirteet voidaan huomioida käytännössä, tekoälyjärjestelmiä ja -sovelluksia kehittävillä tahoilla pitää olla tietoa ja taitoa menetelmistä, joiden avulla ihmisten tarpeisiin voidaan vastata käytännössä. Käyttäjien ja muiden asianosaisten osallistaminen suunnitteluprosessiin on olennainen osa näitä menetelmiä.

Tekoälyjärjestelmiä kehittävien yritysten kyvyt edellä kuvattuun ihmiskeskeiseen suunnitteluun ovat vielä rajallisia. Tekemämme yrityshaastattelututkimus paljastaa, että ainakaan osa yrityksistä ei ota kehitystyössä ihmiskeskeisiä suunnitteluperiaatteita juurikaan huomioon. Tulosten perusteella tunnistimme neljä ihmiskeskeisen tekoälysuunnittelun käytännön päähaastetta:

1. Ihmiskeskeinen suunnittelutyö on eriytynyt teknisestä kehityksestä.
2. Asiakkailla eli tilaajilla on liian keskeinen rooli käyttäjien tarpeiden lähteenä, mikä saattaa johtaa johtaa siihen, että loppukäyttäjien tarpeita ei ymmärretä kunnolla.
3. Tekoälyn mukautuvaan luonteeseen liittyvien käyttöliittymien suunnitteluperiaatteita hyödynnetään puutteellisesti.
4. Tekoälyteknologiaa kehittävissä yrityksissä ei aina ole arvopohjaista tekoälyn ymmärrystä.

Toisaalta yrityksissä on yhä kasvavaa tietoisuutta tekoälyn eettisistä kysymyksistä, ja erityisesti dataan liittyvät yksityisyysongelmat pyritään yleensä ottamaan kattavasti huomioon. Lisäpanostuksia tekoälyn ihmisen huomioivaan näkökulmaan kuitenkin tarvitaan.

10.2 Tekoäly kestävän kaupunkikehityksen muutosvoimana

Kaupungit ja kaupunkiseudut ovat yhteiskunnallisesti keskeisiä toiminnan alueita, joihin tekoälyjärjestelmiä ja sovelluksia on kehitetty ja joita tutkitaan laajasti erilaisissa älykaupunkihankkeissa.[5] Kaupungin kehittymistä ja kaupunkilaisia voivat tukea hyvin erilaiset tekoälypalvelut, jotka muuttavat ihmisten ja infrastruktuurien toimintaa. Älykaupungin keskeisiä toiminta-alueita on kuvattu lyhyesti alla.

Tehokas, miellyttävä ja turvallinen liikkuminen. Tekoälyn avulla voidaan tehostaa julkista liikennettä sekä yksityisautoilua optimoimalla kaupunkilaisten reittejä. Kevyitä, ympäristöystävällisiä liikennevälineitä, kuten yhteiskäyttöisiä polkupyöriä, voidaan

5 Martin, C. J., Evans, J., & Karvonen, A. (2018). *Smart and sustainable? Five tensions in the visions and practices of the smart-sustainable city in Europe and North America.* Technological Forecasting and Social Change, 133, 269–278.

tarjota siellä, missä niitä kulloinkin tarvitaan. Edellä mainittuja tarkoituksia varten voidaan tarjota esimerkiksi sovelluksia, jotka toimivat luonnollista kieltä hyödyntävien käyttöliittymien avulla. Jalankulkijan hyvää kokemusta kaupungista voidaan lisätä ehdottamalla kauniita reittejä, analysoimalla turvallisuusuhkia ja lisäämällä valaistusta tarpeen mukaan. Myös autonomiset kulku- ja kuljetusvälineet, kuten itseajavat autot ja droonit, voivat palvella kaupunkilaisia jo lähitulevaisuudessa.

Hiilineutraalit rakennukset ja asuminen. Sekä julkisten että asuinrakennusten energian käyttöä voidaan mukauttaa älykkäästi sen mukaan, mikä on kunkin rakennuksen tai tilan käyttöaste tai -tarkoitus. Lisäksi rakennusten käyttäjien toimintaa voidaan ohjata energiatehokkaampaan suuntaan ehdottamalla uusia käyttäytymisen muotoja esimerkiksi paikkatunnisteisen mobiilipalvelun välityksellä. Myös asumiseen ja rakentamisen infrastruktuuriin keskeisesti liittyvä jätteenkäsittelyn optimointi on tärkeä osa kestävää kaupunkitoimintaa.

Saavutettavat ja oikeudenmukaiset hyvinvointipalvelut. Kansalaisten hyvinvointiin ja terveydenhuoltoon liittyvien palvelujen järjestämistä voidaan tukea päätöksenteon tukijärjestelmillä. Näissä järjestelmissä tekoäly toimii asiantuntijan "kumppanina" ja auttaa optimoimaan resursseja. Kaupunkilaisen näkökulmasta tavoitteena ovat sujuvat ja aina saatavilla olevat palvelut, joita tarjotaan ihmisille ennakoiden, jopa ennen kuin he itse tiedostavat tarpeensa. Valtiovarainministeriön johtaman *AuroraAI-ohjelman* palveluiden ihmiskeskeinen lähestymistapa on erinomainen esimerkki siitä, miten ihmisten tarpeet ja elämäntapahtumat on otettu tekoälypalveluitten kehityksen keskiöön.

Kaupunkilaisten vuorovaikutteinen osallistuminen kaupunkikehitykseen. Kuntalaisten ja päättäjien aktiivinen vuorovaikutus on keskeinen tekijä hyvän kaupunkikokemuksen kehittämisessä.[6] Tekoälyä hyödyntävät menetelmät voivat houkutella entistä laajemmin kaupunkilaisia esittämään toiveitaan ja ideoitaan kaupungin kehittämiseksi.

Esimerkiksi chatbotteja voidaan integroida helposti sähköisiin palveluihin palautteen keruuta tai kysymysten esittämistä varten. *RoboCivics*-tutkimuksessamme on kokeiltu nuorten osallistumista yhteiskunnalliseen toimintaan sosiaalisten robottien välityksellä sekä kouluissa että muissa nuorille tarkoitetuissa kaupunkitiloissa. Päättäjien palaute kaupunkilaisille on ensisijaisen tärkeää, jotta vuoropuhelu on vaikuttavaa.

Muita kaupunkiin liittyviä tekoälysovelluksia ovat esimerkiksi kaavoituksen kehittäminen tekoälyn avulla sekä vuorovaikutteista älyteknologiaa hyödyntävien kulttuuri- ja viihdepalveluiden tarjoaminen kaupunkitilassa.

Yhteenvetona voidaan todeta, että tekoälyn soveltaminen tarjoaa valtavat mahdollisuudet kaupunkien kestävään kehitykseen sen kaikilla osa-alueilla. Näitä ovat ympäristö ja taloudelliset resurssit sekä sosiaalinen kestävyys, johon sisältyvät ihmisryhmien laaja huomiointi

6 Andreani, S., Kalchschmidt, M., Pinto, R., & Sayegh, A. (2019). *Reframing technologically enhanced urban scenarios: A design research model towards human centred smart cities*. Technological Forecasting and Social Change, 142, 15–25.

ja saavutettavuus. Parhaimmillaan ihmisten arki ja työ sujuvoituvat, resursseja säästyy ja ihmisten elämänlaatu paranee. Uhkina voidaan nähdä etupäässä ihmisten yksityisyyden rikkoutuminen ja siihen liittyvän ”isoveli valvoo” -kulttuurin syntyminen sekä ihmisten yleinen huolestuminen tekoälyn toiminnasta. Tämä saattaa johtaa siihen, että sovellusten tarjoamia mahdollisuuksia vieroksutaan.

10.3 Viestit päättäjille ja kansalaisille: Ihmiskeskeisyys on avaintekijä tekoälyn vaikuttavuudessa

Tekoälyn avulla kaupunkiympäristöt voivat muuttua sujuvammiksi, tehokkaammiksi ja oikeudenmukaisemmiksi. Näin tekoäly voi osaltaan edistää kestävää kehitystä. Vaikuttavuuden saavuttaminen edellyttää, että ihmisiä koskevat uhat ja mahdollisuudet huomioidaan tekoälyjärjestelmien suunnittelun alusta alkaen.

10.3.1 Politiikkasuosituksia päättäjille

Ihmiskeskeisen tekoälyn koulutusta pitää lisätä. Tekoälyn sosioteknisten tekijöiden hallintaa on tärkeää opettaa eri oppialoilla ja monitieteisesti. Yliopistoista ja ammattikorkeakouluista valmistuvat ammattilaiset tarvitsevat osaamista menetelmistä, jotta he pystyvät huomioimaan suunnittelemiensa tekoälytuotteiden pitkäaikaisia sosioteknisiä vaikutuksia. Ymmärrystä erilaisten datalähteiden koostamisesta ja henkilödataan liittyvistä yksityisyyskysymyksistä on nivottava koulutukseen.

Tekoälyjärjestelmien hankintaprosessiin on sisällytettävä selkeät sosiotekniset kriteerit. Hallinnon eri tasojen on huomioitava käyttäjien ja muiden toimijoiden tarpeet, kun hankitaan tekoälyä hyödyntäviä järjestelmiä. Hankintaprosessissa hyvä käyttäjäkokemus ja eettiset tekijät on otettava ensisijaisten valintakriteerien listalle toimittajaa valittaessa ja sopimusta tehtäessä.

Kansalaiset on otettava mukaan tekoälysovellusten suunnitteluun. On tärkeää, että tekoälyjärjestelmien käyttäjät ja muut niiden vaikutuspiirissä olevat ihmiset kertoisivat näkemyksensä sovelluksesta jo suunnittelun varhaisessa vaiheessa siten, että eri ihmisryhmät otetaan huomioon. Kaupunkilaisille suunnattuja tekoälysovelluksia voidaan kokeilla ennen niiden laajamittaista käyttöönottoa esimerkiksi living lab -menetelmillä. Kokeiluista kerätään tietoa ihmiskeskeisen tekoälyn jatkokehityksen tueksi.

Teknisistä järjestelmistä ja sovelluksista ei tule ihmiskeskeisiä itsestään, vaan ihmiskeskeisyyteen täytyy panostaa. Tämä panostus maksaa itsensä takaisin sekä rahallisesti että parantamalla kansalaisten hyvinvointia.

10.3.2 Suosituksia kansalaisille

Älä **pelkää tekoälyä, se on vain kehittynyttä tekniikkaa.** Tekoäly on vain yksi jalostunut tietotekniikan muoto, ei mitään “maagista” tai lähtökohtaisesti epäilyttävää. Kaupungit tai muutkaan arkiympäristömme eivät ole muuttumassa tunnistamattomaksi scifi-näyttämöksi, vaan tekoäly voi tuoda elämiimme sujuvuutta ja turvallisuutta sekä edistää elinympäristömme kestävää kehitystä.

Vaadi parempaa palvelua myös tekoälyltä. Jotta digitaalisten palveluiden käyttö olisi sujuvaa ja turvallista, vaatikaa niiltä parempaa käytettävyyttä, oli niissä tekoälyä tai ei. Älkää koskaan syyttäkö itseänne, jos ette osaa käyttää jotakin järjestelmää tai jos sen toiminnot vaikuttavat epäselviltä – syy on aina huonossa suunnittelussa tai teknisissä ongelmissa. Antakaa palautetta järjestelmän tarjoajalle.

Hanki tietoa tekoälyn toiminnasta. Ymmärrys tekoälyn toiminnan perusteista ja vaikutuksista on kansalaistaito, joka kannattaa hankkia. On tärkeää ymmärtää, miten meistä kerättyä dataa voidaan käyttää hyviin tarkoituksiin. Tekoälyn perusteista on saatavilla verkkokoulutusta, kuten *Elements of AI -tekoälykurssit*.



KUVA 1. Elias-robotti opettaa lapsia koulussa

10.4 Botit auttavat ja kannustavat kaupunkilaisia

Keskusteluroboteilla eli chatboteilla ja sosiaalisilla roboteilla on eräänlaista älykkyyttä, jonka avulla ne voivat toimia ihmisten kanssa vuorovaikutustilanteissa. Tampereella on kokeiltu erilaisia bottiratkaisuja kaupunkilaisten hyvinvoinnin ja osallistumisen edistämiseksi.

Osaamisbottia kokeiltiin Tampereen kaupungin työllisyyspalveluissa, kun kartoitettiin työnhakijoiden osaamista. Osaamisbotti-keskusteluroboti haastatteli työnhakijaa, auttoi tätä tunnistamaan vahvuuksiaan ja osaamistaan, ehdotti erilaisia työpaikkoja sekä motivoi tätä työnhaussa. Osaamisbotti toimi lisäksi omavalmentajan ja työnhakijan viestinnän välittäjänä.



KUVA 2. Pepper-robotti asiakaspalvelupisteellä.

Humanoidirobotti Elias toimii kieltenopettajien apuna Tammelan koululla. Robotin opettamien tuntien teemoina on kielen perusteita. Elias-robotin kanssa voi keskustella teemasta yksinkertaisin lausein. Eliaksen sympaattinen vuorovaikutus ja kannustava opetustyyli vetoavat sekä oppilaisiin että opettajiin. Robotti tekee opetuksesta entistä monipuolisempaa, ja lapset pääsevät tutustumaan robotiikkaan pienestä pitäen.

Tampereen yliopisto teki tutkimuksen palvelurobotti Pepperin käytöstä Tampereen kaupungin palvelupisteessä Frenckellissä. Pepper tarjosi tietoa palvelupisteen toimijoista sekä Tampereen tapahtumista ja palveluliikenteestä. Tutkimustulosten mukaan asiakkaiden kokemukset olivat pääosin myönteisiä. Pepper sai myös aikaan positiivista vuorovaikutusta asiakkaiden välille. Suuri osa asiakkaista ei kuitenkaan hakeutunut vuorovaikutukseen robotin kanssa.

Yhteistä edellä kuvatuille kaupunkilaisten kanssa kokeilluille boteille on se, että ne voivat kannustaa ihmisiä aktiivisuuteen tai vuorovaikutukseen kaupungin työntekijöiden kanssa sekä auttaa oppimisessa ja henkilökohtaisten ratkaisujen löytämisessä. Näin botit voivat lisätä kaupunkilaisten osallistumista, toimijuutta ja hyvinvointia. Koska kaikki kansalaiset eivät kuitenkaan halua olla vuorovaikutuksessa teknologian välityksellä, on tärkeää, että tarjolla on aina vaihtoehto ottaa suoraan yhteyttä ihmiseen.

11 Tekoäly tieteellisessä tutkimuksessa

Kai Puolamäki
apulaisprofessori
Tietojenkäsittelytieteen osasto & Ilmakehätieteiden keskus INAR
Helsingin yliopisto

11.1 Mistä on kyse?

Tässä kirjoituksessa keskityn tekoälyn hyödyntämiseen luonnontieteissä ja muilla mittauksiin ja mallintamiseen perustuvilla tieteenaloilla. Jaan myös näkemykseni tekoälyn mahdollisuuksista lyhyellä ja keskipitkällä aikavälillä.

Näen tekoälyn tieteellisessä tutkimuksessa nimenomaan mahdollisuutena ja yritän avata tätä jäljempänä. Kehitys on kuitenkin huimaavan nopeaa, ja tästä syystä näen suurimpana riskinä sen, että huolimatta Suomen tekoälytutkimuksen ja muun tieteellisen tutkimuksen korkeasta tasosta emme onnistu hyödyntämään kaikkia tekoälyn mahdollisuuksia. Tieteellinen tutkimus mahdollistaa korkeatasoisen opetuksen yliopistoissamme. Useat tieteellisen koulutuksen saaneet luovat uransa ja hyödyntävät oppimaansa yksityisellä tai julkisella sektorilla akateemisen maailman ulkopuolella. Siksi panostus tieteelliseen tutkimukseen hyödyntää suoraan myös yrityksiä ja muuta julkista sektoria, jotka käyttävät tekoälyä hyödykseen.

Tekoälyn hyödyntäminen tieteenalalla voidaan näkemykseni mukaan jakaa karkeasti kolmeen vaiheeseen:

Ensimmäisessä vaiheessa sovelletaan olemassa olevia tekoälymenetelmiä ja laskennallisia käsitteitä sellaisinaan tai pienin muutoksin. Yksi tekoälymenetelmien hyödyistä onkin, että samoja menetelmiä voidaan usein soveltaa eri sovelluskohteisiin. Tietoaineistoja voidaan esimerkiksi analysoida visualisoinneilla tai ryhmittelyillä (klusteroinneilla). Voidaan rakentaa ennustavia malleja, joiden avulla esimerkiksi satelliittikuvista voidaan tunnistaa erilaisia piirteitä.

Toisessa vaiheessa kehitetään uusia menetelmiä ja laskennallisia käsitteitä. Niillä voidaan tutkia ilmiöitä, joita ei ole voitu tutkia aikaisemmin. Tällainen työ vaatii usein hyvin läheistä ja pitkäjänteistä yhteistyötä tekoälytutkijoiden ja muun tieteenalan edustajien kanssa.

Kolmannessa vaiheessa tekoälymenetelmät ja laskennalliset käsitteet ovat erottamaton osa tieteenalan käsitteitä ja menetelmiä – samalla tavalla kuin matematiikan ja tilastotieteen hyvinkin syvälliset käsitteet ovat nykyään erottamaton osa luonnontieteitä ja tieteentekijöiden perusosaamista. Tällöin myös hämärtyy raja sen välillä, mikä on tekoälyä ja mikä ”muuta” tiedettä.

Joillakin tieteenaloilla ollaan vasta aloittelemassa ensimmäistä vaihetta, toisilla ollaan jo pidemmällä, lähellä kolmatta vaihetta. Enää tuskin kuitenkaan löytyy luonnontieteen alan tutkijaa, joka voi jättää tekoälymenetelmät täysin huomioimatta.

Tekoäly mahdollistaa siis parhaimmillaan vastaamisen aivan uudenlaisiin kysymyksiin, joita ei ole aikaisemmin ehkä osattu edes ajatella. Esimerkki pitkälle edenneestä tieteenalasta on bioinformatiikka, joka on syntynyt biotieteiden, matematiikan ja tietojenkäsittelyn sekä tekoälyn yhteistyöstä. Jos kärjistäen aikaisemmin voitiin tutkia yksittäisten geenien toimintaa, nyt tietojenkäsittelyn ja tekoälyn menetelmillä voidaan tutkia vaikkapa eliön kaikkien geenien vuorovaikutuksia yhdellä kertaa. Se ei vanhoilla menetelmillä ollut lainkaan mahdollista. Uusien menetelmien ansiosta biologisia prosesseja voidaan ymmärtää syvällisemmin, minkä seurauksena esimerkiksi sairauksia voidaan hoitaa tehokkaammin.

Erityisen arvokasta tämä tekoälyn tuoma potentiaali on maailmalle nykytilanteessa, jossa on ratkottavana monia vaikeita ongelmia, kuten ilmaston ja ekosysteemin muutokset. Näiden ongelmien ratkaiseminen ja hyvien päätösten tekeminen edellyttävät laajojen järjestelmien ymmärrystä, missä tekoäly on arvokas työkalu.

Tekoäly ja koneoppiminen, joka on tekoälyn mahdollistava teknologia, ovat laajoja ja ei-tarkkarajaisia käsitteitä.

Kummassakin määrittelevänä ominaisuutena on se, että mitä pidemmälle edetään, sitä abstraktimpia ongelmia voidaan ratkaista. Perustehtävä koneoppimisessa on esimerkiksi *luokittelu*, jossa tehtävänä voi olla vaikkapa eritellä satelliittikuvista maaperän tyyppejä. Sen sijaan, että tietokoneeseen ”ohjelmoitaisiin” luokittelusääntöjä, annetaan tietokoneelle esimerkkikuvia, joissa on erilaisia maaperätyyppejä. Näiden perusteella tietokone ”oppii” ratkaisemaan tehtävän. Se siis löytää automaattisesti ehkä hyvinkin monimutkaiset säännöt – tai tekoälykielellä *mallin* – joiden perusteella se voi luokitella kuvat eri maaperätyyppeihin.

Viimeisten vuosikymmenten merkittäviä kehityksiä ovat olleet ohjelmistojen kehitys ja laskentakapasiteetin kasvu. Käytännössä tämä tarkoittaa sitä, että osaava tekijä voi soveltaa edistyneitä tekoäly- ja koneoppimismenetelmiä suhteellisen pienellä vaivalla esimerkiksi edellä kuvatun kaltaisen ongelman ratkaisussa.

Tekoälymenetelmien vahvuus on siis se, että ne pystyvät automaattisesti rakentamaan hyvinkin monimutkaisia säännöstöjä esimerkiksi yllä mainitun tapaisen luokitteluongelman ratkaisemiseksi. Sanotaan, että monet (mutta eivät kaikki) tekoälymenetelmät ovat tästä syystä mustia laatikoita, joiden sisään ei näe. Ei siis ole selvää, millä perusteella ne toimivat.

Tekoälyjärjestelmien musta laatikko -luonne onkin yksi niiden suurimpia haasteita. Tieteessä tärkeää on paitsi tehdä hyviä ennustavia malleja myös ymmärtää näitä malleja. Miten tieteellinen tieto voi lisääntyä, jos mallin toimintaa ei voi ymmärtää? Miten voi luottaa tekoälymallin toimivan oikein varsinkaan uusissa olosuhteissa, jos sen toimintaperiaatteet ovat hämärän peitossa? Tämä ongelma ei kuitenkaan koske vain tiedettä: täsmälleen sama ongelma on tullut vastaan myös automatisoidussa päätöksenteossa.

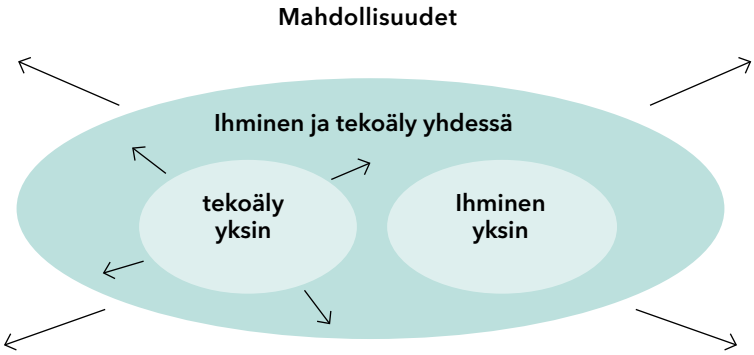
Siinä olisi tärkeää tietää, millä perusteilla tietojärjestelmä kulloinkin päätökseensä päätyy, olipa kyse sitten vaikkapa luottopäätöksen tekemisestä.

Tekoälymenetelmät yleensä ”oppivat” datasta. Ilman dataa ei voi oppia. Tästä syystä avoin data ja helposti saatavat tietoaineistot ovat olleet ja ovat jatkossakin keskeisiä tekoälyjärjestelmien kehitykselle.

11.2 Tekoäly muutosvoimana

Kuten edellä kirjoitin, tämän kirjoituksen kontekstissa näen tekoälyn ensisijaisesti suurena mahdollisuutena. Suurin riski on, että jätämme mahdollisuudet hyödyntämättä. Tekoälyn avulla voidaan hallita yhä suurempia kokonaisuuksia ja vastata haastaviin ja vaikeisiin kysymyksiin, joita ei ehkä ole voitu edes esittää aikaisemmin.

Itse näen yhtenä tekoälyn suurimpana haasteena ja siten myös kehitysmahdollisuutena nykyisten järjestelmien musta laatikko -luonteen. Uskon, että todellinen läpimurto saavutetaan, kun tekoälyn vahvuudet (suurten tietoaineistojen käsittely) voidaan yhdistää ihmisen vahvuuksiin (laaja-alainen asiantuntemus, arvot). Tekoälyn ja ihmisen yhteistyö mahdollistaa paljon enemmän kuin kumpikaan yksin. Tämä on myös asia, jota Suomessa tutkitaan aktiivisesti, allekirjoittanut mukaan lukien.



KUVA 1. Kuvateksti puuttuu! Alt-teksti puuttuu!!

Näen, että monella tieteenalalla olemme menossa kohti edellä mainittua kolmatta vaihetta ja ”virtuaalisia laboratorioita”. Niissä tutkitaan paitsi luontoa ja sen ilmiöitä myös tekoälyn ja laskennallisten simulaatioiden avulla rakennettuja malleja maailmasta. Itse esimerkiksi tutkin Suomen Akatemian rahoittamassa tutkimuksen huippuyksikössä ilmakehätieteisiin liittyvää virtuaalista laboratoriota.[1] Aivan samalla tavalla kuin tosielämän laboratorio tarvitsee mikroskooppinsa ja mittalaitteensa, virtuaalinen laboratorio tarvitsee tekoälytyökalunsa. Niiden avulla voidaan ymmärtää tekoälyyn pohjautuvia malleja ja siten luontoa ja muita tutkittavia ilmiöitä.

1 Virtuaalinen laboratorio ilmakehän molekyylitason reaktioille ja faasimuutoksille, tutkimuksen huippuyksikkö, Suomen Akatemian huippuyksikköohjelma 2022–2029.

Tämä ymmärrys on tärkeää paitsi tieteellisen tiedon lisääntymisen kannalta myös siksi, että voimme vastata tosielämän vaikeisiin kysymyksiin ilmaston ja ekosysteemien muutoksesta sosiaalisiin haasteisiin. Näiden ilmiöiden taustalla olevat prosessit ovat monimutkaisia, ja ongelmien ratkaisu vaatii syvällistä ymmärrystä.

11.3 Viestit päättäjille ja kansalaisille

Asiantuntijoiden kouluttaminen on tärkeää. Erityisesti tarvitsemme huippuasiantuntijoita, jotka ymmärtävät sekä eri tieteenaloja että tekoälyä. Sen lisäksi, että digitaitoja opetetaan laajalti kansalaisille, eri alojen tieteelliseen koulutukseen pitäisi sisällyttää peruskoulutusta tekoälytekniikoista – samalla tavalla kuin esimerkiksi matematiikan ja tietojenkäsittelyn opinnot ovat nykyään osa monen tieteenalan perustaitoja – ja myös syventävää koulutusta tekoälystä pitäisi kehittää. Yliopistossa tutkimus ja opetus ovat sidoksissa toisiinsa. **Tutkimuksen tukeminen** tukee paitsi tieteellisen tiedon tuottamista ja tieteen kehitystä myös sitä, että yksityiselle ja julkiselle sektorille saadaan laaja-alaisia asiantuntijoita, jotka osaavat soveltaa tekoälyä ja voivat hyödyntää uusimpia tutkimustuloksia.

Avoin tiede, avoin data ja avoimet ohjelmistot ovat avainasemassa. Esteitä datan julkaisemiselle tulisi poistaa. Ainakin julkisin varoin tuotetun datan tulisi olla lähtökohtaisesti julkista ja rajoituksitta eri palveluissa hyödynnettävissä, ellei ole erillistä esimerkiksi yksityisyydensuojaan liittyvää syytä olla julkaisematta dataa. Datan lisäksi tarvitaan ohjelmistoja, algoritmeja ja ymmärrystä. Avoimet tekoälyohjelmistot[2] muodostavat perustan tekoälyn käyttämiselle tieteissä. Niitä voisi verrata vaikkapa tieverkostoon, joka mahdollistaa helpon liikkumisen ja kaupankäynnin.

Tekoälystä tulee erottamaton osa tieteenalan käsitteitä ja menetelmiä, kuten matematiikan ja tilastotieteen hyvinkin syvällisistä käsitteistä on tullut erottamaton osa luonnontieteitä ja tieteentekijöiden perusosaamista. Raja sen välillä, mikä on tekoälyä ja mikä ”muuta” tiedettä, hämärtyy. **Virtuaaliset laboratoriot** ovat tulevaisuutta, ja tätä kehitystä on yllä mainituin keinoin syytä tukea.

Tekoälyn monet potentiaaliset vaarat ja ongelmat ovat tuttuja jo ajalta ennen tekoälyä. Ne liittyvät muun muassa yksityisyydensuojaan ja syrjintään. Tietojärjestelmien tekoälykomponentteja ja muita osia ei välttämättä ole helppo erottaa toisistaan, ja siksi on varottava tekemästä sääntelyä erityisesti tekoälylle ilman hyvää perustetta. Jos sääntelyä tarvitaan, **saman sääntelyn pitäisi koskea tietojärjestelmiä yleisesti**. Sääntelyn ei pidä olla sidoksissa siihen, onko tietojärjestelmissä tekoälyä vai ei. Pahimmissa tapauksissa sääntely estää uusien teknologioiden ja niiden käytön kehitystä. Erityisesti on huolehdittava siitä, että sääntely ei tarpeettomasti estä tai vaikeuta tieteellistä tutkimusta.

2 Avoimella ohjelmistolla tarkoitan tässä avoimen lähdekoodin määritelmän mukaisesti lisensoituja ohjelmistoja.

12 Tekoäly syrjäytymisen estämisessä

Peter Ylén,
johtava tutkija VTT
Kehityspäällikkö Espoon kaupunki

12.1 Mistä on kyse?

Nuorten syrjäytyminen on kasvava yhteiskunnallinen ongelma. Tällä hetkellä syrjäytyneiden nuorten määrä on noin 69 000. Kymmenessä vuodessa syrjään jääneiden määrä on kasvanut noin 15 000 nuorella. Sitran ja THL:n tutkijat arvioivat, että inhimillisen kärsimyksen lisäksi nuorten syrjäytyminen maksaa yhteiskunnalle vuodessa noin 1,4 miljardia euroa.[1]

Maahanmuuttajien onnistuneella kotoutumisella ja työllistymisellä on vuosi vuodelta suurempi strateginen merkitys pääkaupunkiseudulle, kun maahanmuuttajien määrä ja väestöosuus kasvavat voimakkaasti. Helsingin vieraskielisen väestön määrän ennustetaan kasvavan lähes 100 000 henkilöllä vuoteen 2035 mennessä. Helsingissä asuisi silloin 196 000 vieraskielistä, mikä olisi 26 prosenttia asukkaista. Pääosa Helsingin väestönkasvusta tulisi vieraskielisistä. Koko Helsingin seudulla vieraskielisiä asuisi 437 000 henkeä, mikä olisi noin 25 prosenttia asukkaista.

Ennusteiden mukaan vieraskielisten määrä Espoossa ja Vantaalla kasvaa suhteellisesti enemmän kuin Helsingissä. Espoossa vieraskielisten kaupunkilaisten määrä on kasvanut yli kaksinkertaiseksi vuoden 2012 jälkeen. Vuonna 2020 se ylitti 55 000 asukkaan rajan. Vuoteen 2035 mennessä 20–64-vuotiaiden vieraskielisten määrän ennustetaan nousevan Espoossa noin 72 000:een, mikä tarkoittaa 131 prosentin kasvua vuoden 2018 alusta. Vuonna 2035 vieraskielisten osuuden Espoon työikäisistä ennustetaan olevan 34,4 prosenttia ja alle 45-vuotiaista työikäisistä 37,4 prosenttia.

Vantaalla ennustetaan, että vuoteen 2030 mennessä ulkomaalaisia tulomuuttajia on vuosittain noin 3 300 henkilöä. Osa ulkomailta tulevista muuttajista on vieraskielisiä, osa kotimaahan palaavia suomen-, ruotsin- tai saamenkielisiä henkilöitä.[2]

Sekä nuorten että maahanmuuttajien työllistyminen ja syrjäytymisen välttäminen ovat yhteiskunnallisesti erittäin merkittäviä asioita. Suorat vaikutukset tuloihin, veroihin ja hyvinvointiin ovat merkittäviä, mutta syrjäytyminen tuo mukanaan myös tarvetta sosiaali- ja terveyspalveluille. Pahimmissa tapauksissa syrjäytymisen myötä yhteisöihin

1 Hilli, P., Ståhl, T., Merikukka, M. & Ristika, T. (2017) **Syrjäytymisen hinta – case investoinninkannattavuuslaskemasta**. Yhteiskuntapolitiikka-lehti 6/2017.
2 Helsingin seudun aluesarjat 2019; Helsingin kaupunginkanslia, Espoon kaupunki ja Vantaan kaupunki 2019.

tulee myös huumeita, rikoksia ja vandalismia. Syrjäytymisellä on toisaalta taipumus periytyä seuraaville sukupolville. Kaikki tämä tulee erittäin kalliiksi yhteiskunnalle. Syrjäytymisessä voidaan panostaa korjaaviin tai ennakoiviin toimenpiteisiin. Korjaavat toimenpiteet tähtäävät siihen, että syrjäytyneet palautetaan ei-syrjäytyneiksi. Ennakoivilla toimenpiteillä pyritään puolestaan varmistamaan jo aikaisessa vaiheessa, että asiat eivät eskaloidu syrjäytymiseksi. Ennakoivat toimenpiteet ovat eettisesti perustellumpia, helpompia ja usein myös edullisempia kumulatiivisesti.

12.2 Tekoäly muutosvoimana

12.2.1 Syrjäytymisen data ja etiikka

Datan ja tekoälyn käytölle on useita rajoitteita syrjäytymisen kontekstissa. Tekoälysovelluksissa on mahdollista käyttää aggregoituja, yhdistettyjä dataklustereita eli datajoukkoja , mutta suurimmat hyödyt saavutetaan käyttämällä yksilötason henkilödataa. Sen käyttöä ohjaavat kuitenkin monet lainsäädännölliset ja eettiset rajoitteet, kuten EU:n yleinen tietosuoja-asetus GDPR sekä laki sosiaali- ja terveystietojen toissijaisesta käytöstä.

Datan ja analytiikan hyödyntämisessä ennakoivilla ja korjaavilla toimenpiteillä on tyypillisesti hyvin erilaiset tavoitteet ja käyttötilanteet. Korjaavat toimenpiteet keskittyvät pääsääntöisesti jo olemassa olevaan ilmiöön, kuten työttömyyteen ja työllistämispalveluihin, ja sen parantamiseen. Data voidaan jopa saada vain yhdestä rekisteristä, joka on rekisterikuvauksessa määritelty tähän tarkoitukseen.

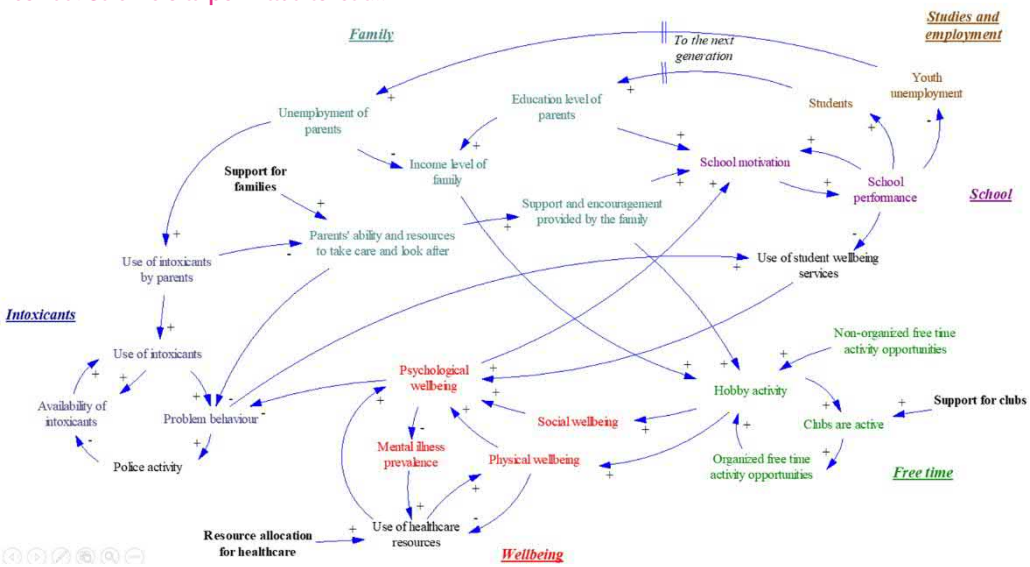
Ennakoivia toimenpiteitä varten tarvitaan dataa usein monista eri rekistereistä, jotta saadaan tietoa kaikista syrjäytymiseen vaikuttavista seikoista. Niitä ovat esimerkiksi terveys, sosiaalipalvelut, työllistyminen, opiskelu ja koulunkäynti monen muun lisäksi. Kaikkia näitä rekistereitä ei ole alun perin tarkoitettu syrjäytymisen estoon. Sen vuoksi eri rekistereistä kerättyä integroitua dataa ei voida käyttää operatiivisiin syrjäytymisen eston palveluihin vaan ainoastaan tutkimustarkoituksiin aggregoituna ja anonymisoituna. Tällöin datan osallistujista on koottu ryhmä, josta henkilöä ei voida enää yksilöidä. Tutkimustuloksia voidaan kyllä hyödyntää palveluiden parantamiseen yleisellä tasolla mutta ei palveluiden kohdentamiseen yksilötasolla.

Tarkastellaan tätä yksinkertaisella esimerkillä. Kuvassa 1 on yksinkertaistettu kaavio nuorten syrjäytymisen syyseuraussuhteista[3]. Jos syrjäytymistä tarkastellaan puhtaasti sosiaali- ja terveyskontekstissa korjaavien toimenpiteiden näkökulmasta (punainen alue kuvan alareunassa), data voidaan kerätä SOTE-lähteistä. Jos taas syrjäytymistä tarkastellaan työllisyyskontekstissa korjaavien toimenpiteiden näkökulmasta (ruskea alue kuvan oikeassa yläreunassa), data voidaan kerätä työllisyyspalveluiden rekistereistä. Jos tahdottaisiin liittää nämä kaksi asiaa yhteen, niin data pitäisi kerätä eri rekistereistä, joita ei ole tarkoitettu eikä rekisterikuvauksessa määritelty tähän yhdistettyyn tarkoitukseen. Tällaista tarkoitusta varten pitäisi perinteisesti perustaa uusi rekisteri, jonka kuvauksessa määritellään käyttötarkoitus ja tarkka prosessinkuvaus ja jolle on valittava rekisterinpitäjä.

---

3 Ylén, Peter; Vainikainen, Sari; Pelkonen, Antti; Suominen, Arho; Mäntylä, Maria; Oksanen, Juha, *Vaikutusten arvioinnin tehostaminen automaattisen tiedonhankinnan ja data-analytiikan avulla*, Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 60/2018.

Löytyykö orginaalia? Tästä ei saa mitään selvää! Saisin sieltä poimittua tekstit...



KUVA 1. Yksinkertaistettu nuorten syrjäytymisen syy-seurauskaavio.

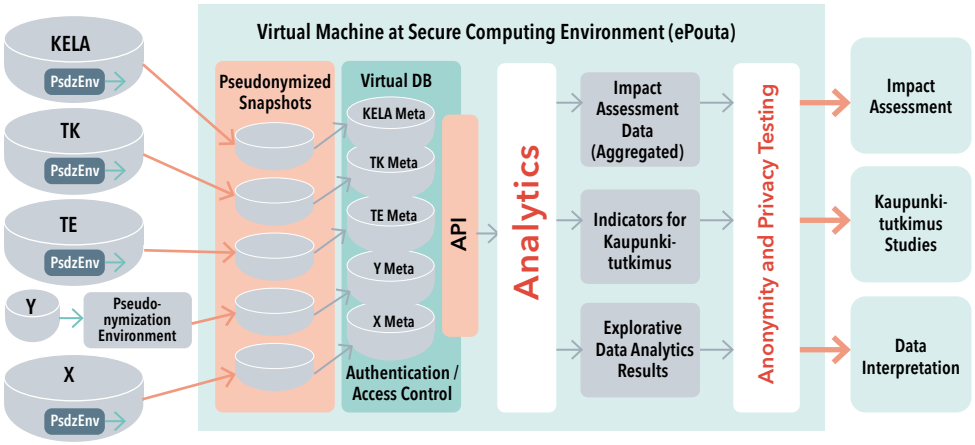
Tilanne monimutkaistuu merkittävästi, kun tarkastellaan syrjäytymisen ennakoivia toimenpiteitä. Syrjäytyminen on hidas prosessi, jonka juurisyyt ja osatekijät saattavat löytyä pitkältä aikajanalta varhaislapsuudesta alkaen. Tällöin data tulisi kerätä kaikista kuvan 1 osa-alueista, kuten perheestä, lapsuudesta, koulunkäynnistä, sosiaalisista suhteista, harrastuksista, häiriökäyttäytymisestä, opiskelusta, työllistymisestä ja sosiaali- ja terveyspalveluista. Näitä asioita koskeva data on hajautettuna ajallisesti ja paikallisesti eri rekistereihin, ja datan yhdistäminen yksilötasolla on erittäin haastavaa.

Vaikka tarkoitus olisi hyvä, eettiset näkökulmat asettavat uusia rajoitteita. Esimerkiksi työllistämistoimista vastaava virkahenkilö saisi yhdistetystä rekisteristä käyttöönsä yksilötasolla työllistettävän henkilön terveyshistorian, joka voi sisältää mielenterveyden hoitotoimia ja muuta arkaluontoista tietoa. Tämä on eettisesti kyseenalaista. Eettisesti astetta kestävämpi on prosessi, jossa virkahenkilö ei pääse käsiksi alkuperäiseen dataan vaan ainoastaan suosituksiin siitä, minkä tyyppiset palvelut tukisivat asiakasta parhaiten. Tämäkin muuttuu eettisesti kyseenalaiseksi, jos virkahenkilö voi suositusten perusteella päätellä asiakkaasta arkaluontoista tietoa, kuten mielenterveyden ongelmia.

Eettisesti kaikkein kestävin ratkaisu on syrjäytymistä koskevan datan käyttö tutkimukseen sosiaali- ja terveystietojen toissijaista käyttöä koskevan lain mukaisesti. Alkuperäisen datan rekisterinpitäjät eivät pääse käsiksi yhdistettyyn dataan. Sen sijaan eri organisaatiossa toimivat tutkijat pääsevät rajoitetusti käsiksi yhdistettyyn pseudonymisoituun dataan. Siitä on poistettu henkilön yksilöinti- ja identifiointitiedot, ja ne on korvattu uudella keinotekoisella tunnuksella. Tämän tunnuksen avulla data voidaan yhdistää yksilötasolla toiseen dataan. Tutkijat soveltavat tekoälyn algoritmeja dataan sekä julkaisevat aggregoidut ja anonymisoidut löydökset tieteellisissä julkaisuissa. Tutkimuksessa tehtyjä löydöksiä voidaan käyttää palveluiden parantamiseen ja kohdistamiseen tilastollisesti. Tutkimusosapuolesta tulee yhdistetyn datan rekisterinpitäjä.

12.2.2 Virtuaaliset tietoaltaat

Virtuaaliset tietoaltaat ovat varteenotettava ratkaisu yhdistää dataa useista rekistereistä. Ne sopivat varsinkin sellaisiin tilanteisiin, joissa yhdistettävä data päivittyy usein. Virtuaalisessa tietoaltaassa data säilyy alkuperäisissä rekistereissä, ja se haetaan vain hetkellisesti analyysiä varten. Analyysin, aggregoinnin, anonymisoinnin ja yksilönsuojan testauksen jälkeen kootut analyysitulokset talletetaan, yhdistetty data tuhotaan analyysialustalta ja alkuperäinen data jää vain rekistereihin, joista se kerättiin yhdistettäväksi. Yhdistetty data on olemassa vain lyhyen analyysin ajan, eikä yhdistettyä yksilödataa talleteta mihinkään. Tämän voidaan tulkita tarkoittavan sitä, että uutta rekisteriä ei muodosteta eikä rekisterinpitäjää tarvita.



KUVA 2. Maahanmuuttajien kotoutumisen ja osaamiskeskuksen virtuaalinen tietoallas.

Kuvassa 2. on esitetty maahanmuuttajien kotoutumisen ja osaamiskeskuksen analytiikkaan suunniteltu virtuaalinen tietoallas[4]. Kyseisessä hankkeessa tutkimusosapuoli eli VTT kerää Espoon kaupungilta sekä eri virastoilta ja sidosryhmiltä henkilödataa. Data toimitetaan pseudonymisoituna virtuaaliselle data-alustalle, jossa se yhdistetään yksilötasolla keinotunnuksen avulla. Yhdistetystä datasta analysoidaan tekoälyä hyödyntäen erilaisia malleja, klustereita ja faktoreita. Tulokset aggregoidaan ja anonymisoidaan, minkä jälkeen testataan vielä yksilönsuojaa ja anonymiteettia. Jos testaus läpäistään, tulokset voidaan viedä jatkojalostukseen tutkijoille ja julkistaa. Yhdistetty yksilödata tuhotaan virtuaaliselta data-alustalta. Tämä prosessi voidaan toistaa automaattisesti aina, kun jokin alkuperäisistä datarekistereistä muuttuu tai päivittyy. Sen ansiosta tulokset ovat aina ajan tasalla.

12.2.3 Tekoälyn mahdollisuudet

Kaikista edellä mainituista rajoituksista ja haasteista huolimatta tekoälyllä on valtava potentiaali syrjäytymisen estämisessä ja hoitamisessa. Tekoälyratkaisut ja analytiikka

4 VTT:n hanketiedote.

yleisemminkin mahdollistavat suurten datamäärien käsittelyn, juurisyiden määrittelyn, toimenpiteiden vaikutusten arvioinnin ja useiden tekijöiden yhteisvaikutusten määrittämisen hyvin heterogeenisesta populaatiosta. Oikein sovellettuna tekoäly voi antaa työkaluja yhteiskunnallisesti merkittävän ongelman ratkaisuun.

On kuitenkin muutamia seikkoja, jotka on otettava huomioon, kun sovelletaan tekoälyä syrjäytymisen ymmärtämiseen ja sen ratkaisemiseen tähtääviin toimenpiteisiin. On ymmärrettävä, että sekä syrjäytyneet nuoret että erityisesti syrjäytyneet maahanmuuttajat ovat todella fragmentoitunut joukko. Siihen kuuluvilla henkilöillä on hyvin erilaisia syrjäytymisen mekanismeja, ja heidän auttamisekseen tarvitaan hyvin erilasia toimenpiteitä. Näin ollen tarvitaan edustava ja runsas aineisto, jonka avulla voidaan muodostaa erilaisia segmenttejä ja faktoreita selittämään tätä heterogeenista ryhmää.

Vastaavasti on otettava huomioon, ettei monista syrjäytymiseen merkittävästi vaikuttavista tekijöistä ole formaalia dataa. Kun olennaiset selittävät tekijät jäävät analyysistä pois, saatu tulos kertoo vain rajallisen osatotuuden. Tuloksiin on suhtauduttava kriittisesti. Kvantitatiivisten analytiikkatulosten integrointi kvalitatiivisiin menetelmiin antaa usein parhaan tuloksen. Syrjäytymisen aikajänne voi olla hyvin pitkä, ja monet ajallisesti kaukana menneisyydessä olevat seikat vaikuttavat syrjäytymiseen. Jos ilmiöön halutaan vaikuttaa ennakoivilla toimenpiteillä, dataa on kerättävä hyvin pitkältä ajalta.

12.3 Toimenpide-ehdotukset

Avoimen datan laatuun ja yhdistettävyyteen on panostettava. Erilaiset alustaratkaisut voivat olla tässä hyödyllisiä. Erilaisten datojen fuusiota (julkinen, yksityinen ja omadata) ja toissijaista käyttöä on edistettävä lakien ja asetusten avulla henkilöiden tietoturvasta tinkimättä. Datan toissijaisen käytön mahdollistamista on harkittava eri hallinnonalojen sisällä ja välillä. Tarkkojen ja ajantasaisten yksilökohtaisten tietojen hyödyntäminen avaisi uusia mahdollisuuksia vaikutusten arviointiin yksityisyydensuojaa kunnioittaen ja GDPR-asetuksen mukaisesti.

Julkisella sektorilla on panostettava tekoälyn osaamiseen menetelmätasolla analytiikan erikoisasiantuntijoiden osalta sekä eri menetelmien mahdollisuuksien ja rajoitusten ymmärtämiseen strategisen johdon osalta. Lisäksi julkisen sektorin eri toimijoiden ja rekisterinpitäjien pitäisi tehdä yhteistyötä holististen ja systeemisten ongelmien ymmärtämiseksi ja ratkaisemiseksi. Ainoastaan yhtä toimijaa ja yhtä datalähdettä hyödyntämällä voi löytää aina vain osaratkaisun syrjäytymiseen.

Lupaprosesseja, jotka koskevat datan keräämistä eri lähteistä ja datan yhdistämistä, on nopeutettava. Laajojen data-aineistojen luvituksessa voi vierähtää yli vuosi kalenteriaikaa. Usein luvitusprosessi on suunniteltu määräaikaiseen datan käyttöön, kun taas tekoälyn hyödyntäminen jatkuvatoimiseen operatiiviseen toimintaan edellyttäisi toistaiseksi voimassa olevaa luvitusta.

Datarekisterien pitäjien olisi päivitettävä rekisterien käyttötarkoitusten kuvauksia, jotta rekisterin sisältämää dataa voitaisiin yhdistää muihin datoihin myös operatiivista käyttöä, ei ainoastaan tutkimuskäyttöä, varten.

13 Tekoälystä ymmärrettävää

Teemu Roos
tietojenkäsittelytieteen professori ?
Helsingin yliopisto

13.1 Tekoälyn lukutaito eli mitä jokaisen tulisi tietää tekoälystä

Euroopan komission tiedonannossa *Tekoäly Euroopassa* (COM/2018/237) tekoäly (engl. *artificial intelligence*, *AI*) määritellään seuraavalla tavalla:

> ”Tekoäly on järjestelmä, joka pyrkii älykkäästi saavuttamaan asetettuja tavoitteita analysoimalla ympäristöään ja toimimalla osittain itsenäisesti.”

Määritelmän jokainen sana on huolella valittu[1], ja määritelmä kiinnittää huomiota tekoälyyn teknologiana, joka mahdollistaa jossain määrin autonomiset toiminnot. Se myös korostaa hyvin määriteltyjen tavoitteiden ja järjestelmän kehittäjien asettamien päämäärien merkitystä.

Edellä mainittu määritelmäkään ei kuitenkaan ole riittävän tarkkarajainen, jotta sitä voitaisiin yksikäsitteisesti soveltaa lainsäädännössä. Onko esimerkiksi lämpöpatterin termostaatti tekoälyä? Termostaatti analysoi ympäristöään (mittaa ympäristön lämpötilaa) ja suorittaa toimintoja (säätää patterin lämpötilaa) ja siten saavuttaa autonomisesti hyvin määritellyn tavoitteensa eli pitää huoneen lämpötilan tasaisena.

Tekoälyn määritteleminen onkin kaikille tekoälyn parissa työskenteleville tuttu kompastuskivi, joka kannattaa kiertää kaukaa, jos suinkin mahdollista. Oli määritelmä mikä hyvänsä, tekoälyn yhteiskunnalliset ja taloudelliset vaikutukset ovat nyt ja tulevaisuudessa valtavat.

Varsinaisten tekoälytuotteiden eli tekoälyohjelmistojen ja -laitteistojen myymiseen liittyvän liiketoiminnan on arvioitu kasvavan vuosina 2021–2024 maailmanlaajuisesti yli 60 prosenttia eli yli 550 miljardiin dollariin vuodessa.[2] Tätä lukua huomattavan paljon muilla aloilla, kuten markkinoinnissa.[3] Kaiken kaikkiaan tekoälyn on arvioitu tuottavan

1 Tosin siitä, miten suomennoksessa tekoäly-termillä viitataan yksikössä "järjestelmään", saattaa tulla kuva, että tekoäly on jokin yksittäinen järjestelmä. Itse pidän parempana puhetapaa, jossa yksittäisiin järjestelmiin viitataan aina termillä tekoälyjärjestelmä ja tekoäly-sanalla viitataan järjestelmiä tutkivaan tieteenalaan tai teknologiaan yleisesti. Englanninkielinen muotoilu "Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals" on tässä mielessä selkeämpi.

2 Statista, "*Artificial Intelligence market size worldwide 2020–2024*", 27.9.2021, kirjoittaja S. Liu, haettu 21.10.2021.

3 S. Hill, P.C. Evans, D. Ryan, N. Zaman & A. Congalton, "*AI transforming the enterprise, Eight key trends in the adoption of artificial intelligence*", KPMG, 2019.

suurempia rahamääriä liikkuu liiketoiminnassa, jossa sovelletaan tekoälymenetelmiä vuosina 2017–2030 yli 15 biljoonan dollarin kasvun maailmantalouteen.[4] Kansantaloudellisesti kaikkien maiden yhteenlasketun bruttokansantuotteen on arvioitu kasvavan tekoälyn johdosta joka vuosi noin 1,2 prosenttia.

Suomi on tekoälyn kärkimaita.[5] Suomessa oli vuonna 2020 Suomen tekoälykiihdyttämön (FAIA) *State of AI in Finland* -raportin mukaan yli 1 250 tekoälyä hyödyntävää yritystä, ja uusia tekoälyä käyttäviä startup-yrityksiä syntyy noin 50 kappaleen vuosivauhdilla.

Taloudellisten vaikutusten ohella tekoälyllä on myös laajoja yhteiskunnallisia vaikutuksia. Nämä vaikutukset ulottuvat jokaiseen kansalaiseen. Algorithm Watch -järjestön raportin mukaan automaattisen päätöksenteon sovellukset lisääntyvät kiihtyvää vauhtia mutta niihin liittyvä demokraattinen keskustelu on yhä puutteellista ja lainsäädäntö laahaa perässä.[7] Raportti suosittaa ”algoritmien lukutaidon” (engl. *algorithmic literacy*) edistämistä, jotta aiheesta käytävä keskustelu on mahdollista kaikille, ei pelkästään teknologia-alan asiantuntijoille.

Tekoälyn perusperiaatteiden osaaminen on tärkeä kansalaistaito:

1. Jokainen kansalainen voi hyötyä tekoälyyn perustuvista tuotteista ja palveluista paremmin, jos tuntee niiden toimintaperiaatteet, mahdollisuudet ja rajoitteet. Tällöin voi myös välttää tekoälyn haitallisia vaikutuksia, kuten yksityisyyden suojan heikentymistä ja eriarvoistumista.
2. Tekoäly tarjoaa merkittäviä taloudellisia mahdollisuuksia kaikilla toimialoilla. Kaikkien ei tarvitse ryhtyä koodareiksi, vaan tekoälyyn liittyviä innovaatioita on mahdollista ideoida, kunhan tuntee yleiset periaatteet.
3. Yksilötason hyötyjen ohella yhteiskunnassa vallitseva laaja tekoälyn lukutaito mahdollistaa faktapohjaisen, valistuneen kansalaiskeskustelun, joka on olennainen osa demokraattista päätöksentekoa. Tekoälyn pelisääntöjä koskevaa päätöksentekoa ei pidä jättää pelkästään harvojen teknologian asiantuntijoiden käsiin.

Peruskoulun opetussuunnitelmassa, toisen asteen koulutuksessa ja monien alojen jatkokoulutuksessa korostetaan nykyään – vaikkei välttämättä suoraan tekoälyä – ainakin tekoälyyn liittyviä tietoja ja taitoja, kuten ohjelmoinnillista ajattelua, teknologiakasvatusta ja medialukutaitoa.[8] Alan koulutusmäärät ovat kasvussa.[9] Tästä huolimatta suurin osa väestöstä tulee vielä pitkään olemaan vailla tekoälyn perustaitoja, ellei osaamisen

4 A.S. Rao ja G. Verweij, "*Global Artificial Intelligence Study: Exploiting the AI Revolution*", PwC, 2017.
5 Tortoise Media, "*The Global AI Index*", haettu 21.20.2021; Suomi on vertailussa sijalla 12.
6 Suomen tekoälykiihdyttämö FAIA, "*State of AI in Finland*", 2020, haettu 21.10.2021.
7 AlgorithmWatch, "*Automating Society Report*", 2020, haettu 21.10.2021.
8 Opetushallitus, "*Perusopetuksen opetussuunnitelman perusteet*", 2014.
9 H. Ailisto, A. Neuvonen, H. Nyman, M. Halén ja T. Seppälä, "*Tekoälyn kokonaiskuva ja kansallinen osaamiskartoitus – loppuraportti*", Valtioneuvoston kanslia, 15.1.2019.

päivittämiseen kohdisteta merkittäviä jatkuvan oppimisen toimia, kuten avoimia yliopisto-opintoja, verkkokursseja[10] ja pieniä osaamiskokonaisuuksia eli niin sanottuja mikrokredentiaaleja.[11]

13.2 Tekoälyn poliittiset ulottuvuudet

Tekoälyä koskevaa sääntelyä ja standardointia kehitetään aktiivisesti sekä kansallisella että kansainvälisellä tasolla. Sääntelyn tarkoitus on yhteiskunnallisten vaikutusten suuntaaminen siten, että haittavaikutukset pystytään minimoimaan ja että ne ovat muun muassa ihmisoikeuksien, tasa-arvon ja yhdenvertaisuuden kannalta hyväksyttäviä. Samalla pyritään maksimoimaan taloudelliset hyödyt turvaamalla yrityksille ennustettava ja avointa kilpailua tukeva toimintaympäristö. Suuri osa sääntelystä liittyy sektorikohtaiseen lainsäädäntöön, esimerkkeinä yhdenvertaisuuslaki, tuote- ja liikenneturvallisuuslait sekä sananvapaus- ja tekijänoikeuslait. Sektorikohtainen lainsäädäntö käsittelee tekoälyn sovelluksia lähtökohtaisesti samalla tavalla kuin muita teknologioita, mikä sopii yhteen EU:n lainsäädännön periaatteena olevan teknologianeutraaliuden kanssa. Toisaalta EU:n ehdottama tekoälyasetus (COM/2021/206) poikkeaa tästä ja tavoittelee nimenomaan tekoälyjärjestelmiä koskevaa lainsäädäntökehikkoa.

Kuten olen valiokuntakuulemisissa tuonut ilmi, olen sitä mieltä, että erikseen tekoälyä koskeva asetus aiheuttaa merkittäviä hankaluuksia. Tämä johtuu tekoälyn määrittelyn vaikeudesta. Hankaluudet määrittelyssä voivat käytännössä johtaa riskiin, että sääntelyä pyritään kiertämään soveltamalla menetelmiä, jotka valitun määritelmän mukaan eivät ole tekoälyä. Lisäksi erillinen tekoälyasetus aiheuttaa väistämättä päällekkäisyyttä sektorikohtaisen lainsäädännön kanssa. Tämä edellyttää huolellista ja työlästä lainsäädäntökehikkojen yhteensovittamista, jotta ristiriitaisuuksilta vältytään. Maito on kuitenkin jo mitä luultavimmin läikkynyt, ja komissio on tuskin enää halukas perääntymään erillisestä tekoälyasetuksesta.

Kun käsitellään asetusehdotusta ja siihen liittyviä valvovien viranomaisten tekemiä linjauksia, jotka tulevat tarkemmin määrittämään asetuksen tulkinnan, on joka tapauksessa tärkeää kuulla sektorikohtaisesti kunkin alan lainsäädännön asiantuntijoita. Muutoin on mahdollista, että tekoälyasetuksen takia päädytään ottamaan joillakin sektoreilla askelia taaksepäin. Tämä sektorikohtaisen asiantuntemuksen ja tekoälyasiantuntemuksen yhteensovittaminen vaatii nimenomaan tekoälyn perusperiaatteiden opettelua kaikilta työhön osallistuvilta eri sektorien asiantuntijoilta. Vastaavaa perehtymistä kulloinkin käsiteltävän sektorin perusperiaatteisiin ja olemassa olevaan lainsäädäntöön vaaditaan myös työhön osallistuvilta tekoälyasiantuntijoilta.

---

10 Helsingin yliopiston ja Reaktorin vuonna 2018 julkaisemalle **Elements of AI -verkkokurssille** oli lokakuuhun 2021 menneessä ilmoittautunut jo noin 120 000 suomalaista ja maailmanlaajuisesti yli 760 000 käyttäjää. Kurssi laajeni osana Suomen EU-puheenjohtajakauden ohjelmaa työ- ja elinkeinoministeriön, EU-komission ja Teknologiateollisuuden 100-vuotissäätiön tuella kaikkiin EU-maihin. Helsingin yliopisto on julkaissut myös tekoälyn etiikkaa käsittelevän **Ethics of AI -verkkokurssin**.

11 M. Mäkelä & K. Mäki, "***Jatkuva oppiminen haastaa osaamiskeskeisyyteen***", Haaga-Helia ammattikorkeakoulu, 22.4.2021, haettu 21.10.2021.

Koska tekoälyyn liittyy taloudellisia intressejä, kohdistuu sääntelyyn merkittävää mielenkiintoa. Tämän vuoksi päätöksenteossa kuultavien asiantuntijoiden valintaan on syytä kiinnittää erityistä huomiota. Minkään osapuolen, kuten yksittäisten yritysten tai yritysten edunvalvojien, kantaa ei luonnollisesti pidä jättää kuulematta tai huomioimatta, mutta on kiinnitettävä huomiota siihen, että kuultavien asiantuntijoiden joukossa on myös riippumattomia asiantuntijoita, kuten yliopistojen ja tutkimuslaitosten tutkijoita laaja-alaisesti useilta tieteenaloilta.

13.3 Tekoälyn hyödyntäminen poliittisessa päätöksenteossa

Edellä mainitun tekoälyä koskevan päätöksenteon lisäksi on aiheellista mainita tekoälyn hyödyntäminen päätöksenteossa – myös poliittisessa. Huolimatta siitä, että jotkut ovat esittäneet utopistisia (tai dystooppisia) ehdotuksia kohottaa tekoäly poliittisten päätöksentekijöiden[12] tai näille tietoa tuottavien asiantuntijoiden[13, 14] asemaan, ei tekoälyn hyödyntäminen päätöksenteossa tosiasiassa poikkea luonteeltaan merkittävästi tavanomaisesta tietopohjaisesta päätöksenteosta: kyse on systemaattisesta tiedon keräämisestä, kerätyn tiedon analysoinnista ja sen pohjalta tuotetun tiedon esittämisestä tavalla, joka tukee päätöksentekoa.

Tekoäly voi toimia päätöksenteon tukena muun muassa silloin, kun haetaan relevanttia tietoa laajoista aineistoista ja prosessoidaan tekstimuotoista dataa esimerkiksi kansalaisilta kerätyn aineiston tiivistämiseksi. Tekoälyn avulla voidaan myös tuottaa niin sanottuja oppivia malleja, joilla ennustetaan esimerkiksi talouden tunnuslukuja tai kansalaisten terveyttä uhkaavien tartuntatautien leviämistä. Tällaisia tekoälymenetelmiä sovelletaan aktiivisesti niin ekonometriassa ja epidemiologiassa kuin muillakin aloilla, ja niillä tuotettu tieto on jatkossa yhä enemmän päätöksentekijöiden tukena. Taustalla olevat työkalut ovat pitkälti samoja kuin tilastotieteessä, kuten regressioanalyysi ja optimointimenetelmät. Tekoäly ja koneoppiminen kuitenkin laajentavat näiden työkalujen soveltuvuutta erityisesti tilanteissa, joissa hyödynnettävä data on monimuotoista ja peräisin useasta eri lähteestä.

13.4 Yhteenveto ja politiikkasuosituksia

Yhteenvetona voidaan todeta, että tekoäly on poikkeuksellisen monitahoinen ilmiö ja se vaikuttaa laajalti kaikilla yhteiskunnan aloilla. Tekoälyn perusperiaatteiden tunteminen on tärkeä kansalaistaito, johon on syytä kansallisesti kohdentaa jatkuvan oppimisen toimia – unohtamatta meistä jokaisen henkilökohtaisia panostuksia oman osaamisemme päivittämiseen.

---

12 O. Jonsson, C.L. de Tena, D. Rubio, C.L. Anadon, I. Menéndez, C.B. Frey & C. Colclough, "*European Tech Insights*", The Center for the Governance of Change, 2021, haettu 21.10.2021.

13 Tulevaisuusvaliokunta, "*Tulevaisuusvaliokunta kuuli tekoälyä tiettävästi ensimmäisenä valiokuntana maailmassa*", tiedote, 9.4.2021, haettu 22.10.2021.

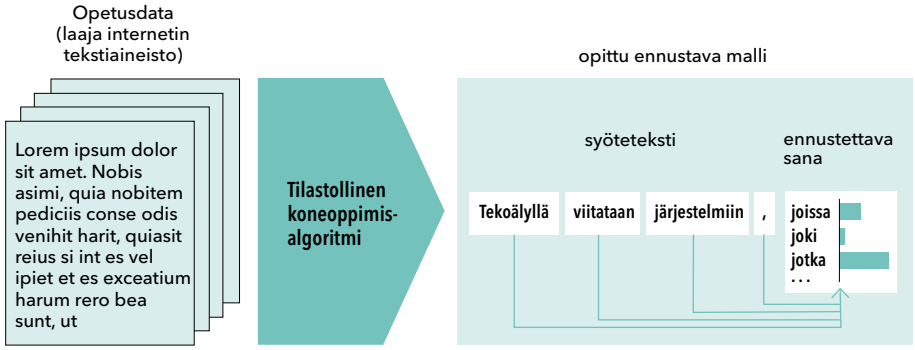
14 F. Fitsilis, "*Artificial Intelligence (AI) in parliaments – preliminary analysis of the Eduskunta experiment*", Journal of Legislative Studies, 10.9.2021.

Tekoälyä käsitellään usein etupäässä teknologisena ilmiönä. Aiheeseen liittyvien taloudellisten ja yhteiskunnallisten intressien yhteensovittaminen ei kuitenkaan ole teknologinen vaan poliittinen kysymys. Tekoälyä koskevassa poliittisessa päätöksenteossa on keskeistä kuulla laajalti asiantuntijoita, etujärjestöjen edustajia ja kansalaisia. Asiantuntijoina esiintyviä kuultaessa on tärkeää tiedostaa, että aiheeseen liittyy merkittäviä taloudellisia intressejä. Tieteellistä asiantuntemusta ei tule sivuuttaa – myös tässä yhteydessä on edelleen syytä painottaa, että kyseessä ei ole pelkästään teknologinen tai luonnontieteellinen vaan myös muun muassa lainopillinen ja yhteiskuntatieteellinen ilmiö.

Suomella on hyvät edellytykset olla jatkossakin tekoälyn edelläkävijämaa. Vaikuttamalla EU-tasolla ja kansainvälisissä verkostoissa Suomi voi saavuttaa huomattavasti kokoaan suuremman painoarvon. Erityisesti standardointitoimintaan on syytä panostaa nykyistä enemmän. Taloudellisten näkökulmien ohella Suomi voi merkittävällä tavalla edistää ihmislähtöisen, kestävän ja kaikkien kansalaisten etua palvelevan tekoälyn kehitystä ja käyttöönottoa.

13.5 Miksi tekoäly ei ole asiantuntija?

Tekoälyn avulla voidaan jo nyt tuottaa pintapuolisessa tarkastelussa aidolta vaikuttavaa kuvaa, videota, ääntä ja tekstiä. Laajoihin tekstimassoihin perustuvat kielimallit, kuten GPT-3, ovat herättäneet paljon mielenkiintoa niin tekoälytutkijoiden kuin kaupallisten toimijoidenkin keskuudessa (ks. kuva 1). Niiden on väitetty jopa pystyvän tuottamaan tieteellisiä artikkeleita tai toimimaan eri alojen asiantuntijoina.



KUVA 1. Ennustavan tekstimallin tuottaminen opetusdatasta. Malli opitaan laajasta tekstiaineistosta tilastollisen koneoppimisalgoritmin avulla. Malli tuottaa todennäköisyysjakauman syötetekstiä seuraavalle sanalle. Mallissa voi tuottaa tekstiä valitsemalla seuraava sana todennäköisyysjakauman mukaisesti, liittämällä se syötetekstin jatkoksi ja toistamalla prosessia, kunnes on saatu haluttu määrä tekstiä. Kuvan esimerkissä, malli valitsisi todennäköisimmin syötetekstin "Tekoälyllä viitataan järjestelmiin," jatkoksi sanan "joissa" tai "jotka".

Tekstimalleihin liittyy kuitenkin merkittäviä haasteita, joista useimmat juontavat juurensa kahdesta perusongelmasta. Ensinnäkään mallit eivät sisällä mitään muuta tietosisältöä kuin sanojen tilastollisia esiintymistodennäköisyyksiä muiden sanojen yhteydessä. Tämä ilmenee usein loogisina aukkoina ja ristiriitaisuuksina tuotetussa tekstissä. Samasta syystä mallin tuottamat väitteet ovat usein virheellisiä. Lisäksi niille on hankala tehdä faktantarkistusta.

Toinen, väitteiden paikkansapitävyydestä riippumaton ongelma on se, että internetistä ladattuja, valtavan suuria opetusaineistoja ei tavallisesti ole ”kuratoitu” riittävän huolellisesti ja ne sisältävät muun muassa selvästi rasistisia tai muuten syrjiviä elementtejä.[15] Näiden puutteiden vuoksi GPT-3:n kaltaisten tekstimallien soveltamista poliittisen päätöksenteon tukena voidaan pitää kyseenalaisena.

15 E. Strickland, "*OpenAI's GPT-3 Speaks! (Kindly Disregard Toxic Language)*", IEEE Spectrum, 1.2.2021, ladattu 22.10.2021.

14 Tekoälystä vastuullista

Jaana Leikas
dosentti, johtava tutkija
Teknologian tutkimuskeskus VTT Oy

14.1 Mistä on kyse?

Algoritmisen päätöksenteon mahdollisuuksien ja eettisten haasteiden ymmärtäminen ja tekoälyn vastuullinen hyödyntäminen ovat aikamme avainkysymyksiä. Taloudellisten hyötyjen ja lisääntyvän tehokkuuden ohella tekoälystä on povattu merkittävää tekijää hyvän yhteiskunnan ja YK:n sosiaalisten kehitystavoitteiden edistämisessä. Julkisessa puheessa tekoälyn avulla pyritään luomaan paitsi vaurautta ja kilpailukykyä myös oikeudenmukaista ja kestävää yhteiskuntaa sekä poliittista ja sosiaalista tasa-arvoa. Parhaassa tapauksessa tekoäly on hyvä työkalu suuren datamäärän ja kompleksisten suhteiden jäsentämisessä. Se voi olla loistava apuri kansalaisten elämän ja organisaatioiden toiminnan suuntaamisessa. Itse asiassa näiden tekoälyyn kohdistettujen odotusten vuoksi sen alikäyttöä ja potentiaalin hyödyntämättä jättämistä voidaan pitää eettisesti arveluttavana ja jopa riskinä yhteiskunnalle.

Kolikolla on myös kääntöpuolensa. Tekoälyn suunnitteluun, kehittämiseen ja käyttöönottoon liittyy ristiriitaisia kysymyksiä. Riskinä voidaan nähdä tekoälyn väärinkäyttö ja sen hyödyntäminen tarkoituksiin, jotka eivät tue ihmiskunnan kukoistusta pitkällä aikavälillä. Tekoälyjärjestelmien kehittäminen voidaan nähdä myös poliittisena kysymyksenä, joka liittyy vallan ja legitiimiyden kysymyksiin. Huolenaiheena on, että tekoälyjärjestelmiä soveltamalla voidaan entisestään lisätä vallan ja vaurauden epäsymmetriaa. Tämä syventäisi olemassa olevaa ja jopa loisi uudenlaista eriarvoisuutta.

14.1.1 Tekoälyn eettiset kysymykset ovat moniulotteisia

Tekoälyn eettiset haasteet, pois lukien superälykästä tekoälyä koskevat kysymykset, voidaan karkeasti jakaa **koneoppimisesta ja algoritmeista juontuviin haasteisiin ja yhteiskunnallisiin haasteisiin.** Koneoppimisesta ja algoritmeista juontuvia haasteita ovat esimerkiksi datan vääristyneisyys, syrjintä ja datan turvallisuus. Yhteiskunnallisia haasteita ovat esimerkiksi muutokset ihmisen autonomiassa sekä ihmisen ja koneen välisessä vuorovaikutuksessa, järjestelmien sosiaalinen hyväksyntä, tasa-arvokysymykset ja yhteiskunnalliset valtarakenteet sekä työn muutokseen ja ekologiseen kestävyyteen liittyvät kysymykset.[1]

1 Esim. Stahl, Bernt Carsten (2021). *Artificial Intelligence for a Better Future*. Springer briefs in research and innovation governance.

Kaikkia tekoälyn negatiivisia vaikutuksia ei luonnollisesti osata arvioida ennalta. Riskien arviointia helpottanee tarkastelun aikajanan asettaminen nykyisyyteen, lähitulevaisuuteen ja pitkälle aikavälille: Millaisia ovat ne tekoälyn riskit, jotka voivat olla nähtävissä jo tänään? Millaisia riskejä voimme kohdata tulevaisuudessa nykyisten teknologioiden yhteydessä? Mitä riskejä on kohdattavissa pitkällä aikavälillä, kun teknologia kehittyy? Näiden kysymysten operationalisointi ja istuttaminen tekoälyn konteksteihin vaatii analyyttisyyttä ja uusia työkaluja.

Ihmiskeskeisyyttä pidetään keskeisenä tekijänä tekoälyn kehittämisessä. Useat Euroopan komission strategiat ja ohjeet korostavat tätä keskeisenä tavoitteena: ihmisten tulee voida luottaa siihen, että tekoälyä käytetään turvallisesti ja lainmukaisesti ja että kansalaisten perusoikeuksia kunnioitetaan. Asiakirjoissa viitataan kuitenkin ihmiskeskeisyyteen vain pinnallisesti keskittymällä lähinnä ihmisoikeuksien toteutumiseen. Sama näkyy teknologian suunnitteluohjeissa: vaikka ihmiskeskeisyyden soveltamisella on pitkät perinteet ihmisen ja teknologian välisen vuorovaikutuksen suunnittelussa, suunnittelun kentältä puuttuu edelleen etiikkaa ja sosiaalista kestävyyttä tarkastelevia teknologian suunnittelumenetelmiä.

14.1.2 Eettiset koodistot ja ohjeet viitoittavat suuntaa

Tilannetta helpottamaan on luotu erilaisia eettisiä koodistoja. Niiden tavoitteena on muun muassa turvata teknologian käytön kohteena olevien henkilöiden asema ja oikeudet. Useiden valtioiden, suuryritysten ja kansainvälisten järjestöjen dokumenteissa, jotka linjaavat tekoälyn etiikkaa, mainitaan eettisinä periaatteina läpinäkyvyys, oikeudenmukaisuus, vahingonteon välttäminen, vastuullisuus ja yksityisyys. Euroopan unionissa vaatimuksina ovat lisäksi tekoälytoiminnan selitettävyys ja ihmislähtöisyys. Eurooppalaisten maiden poliittisissa dokumenteissa vaaditaan lisää luottamusta, vastuullisuutta, vastuuvelvollisuutta, läpinäkyvyyttä ja turvallisuutta tekoälyn kehitykseen ja hyödyntämiseen.

Myös tutkimuskenttä on kehittynyt nopeasti, ja **kehitteillä on lukuisia erilaisia ohjeistuksia ja käytäntöjä sosiaalisesti hyväksyttävän tekoälyn suunnitteluun**. Tutkijat ovat niin ikään esittäneet näkemyksiä eettisten periaatteiden juurruttamiseksi tekoälyn ohjaus- ja koordinaatiomekanismeihin, joista osa on kirjattu soveltavasti osaksi EU:n tekoälyn hallinnan toimenpidesuosituksia[2] sekä OECD:n periaatteiksi[3].

Tekoälyä ohjaavia eettisiä suosituksia, ohjeistuksia ja periaatteita on kritisoitu siitä, että niiden käytännön vaikutus on vähäinen. **Ohjeiden noudattaminen on suunnittelijoille ja kehittäjille hankalaa, sillä ohjeet ovat tulkinnanvaraisia eivätkä ne tarjoile yhtä ainoaa tekoälyn eettisyyttä määrittävää periaatetta tai riittävää mittaria**. Ne eivät myöskään näytä juurtuvan politiikkatoimenpiteisiin. Lähinnä ne mahdollistavat niin sanotun etiikkapesun yrityksissä.

---

2 AI HLEG (2019). Ethics Guidelines for Trustworthy AI.
High-Level Expert Group on Artificial Intelligence, European Commission, April 2019.

3 OECD (2017). Recommendation of the Council on Open Government. Paris: OECD.

Eettisiä periaatteita voidaan käyttää kompassina, kun etsitään tapoja ymmärtää tekoälyn etiikkaa. Periaatteiden nimeäminen helpottaa asioiden punnintaa, mutta ne eivät aina vie tekoälyajattelua kovin pitkälle. Suunnittelijalle voi tulla eteen tilanteita, joissa periaatteet ovat ristiriidassa keskenään. Yksi syy tähän on se, että **teknologian etiikkaa koskevat kysymykset ovat aina tilannesidonnaisia.** Teknologian vaikutukset eivät koske pelkästään suoria käyttötilanteita, vaan ne ovat käyttötilanteen kautta sidoksissa erilaisiin sidosryhmiin, joilla saattaa olla ristiriitaisia tavoitteita. Eri vaakakupeissa voivat olla myös yhteiskunnan etu ja yksilön etu. Samoin yksilön oikeus päätöksentekoon voi sotia jonkun toisen etua vastaan. Siksi etiikan tarkastelu vaatii laajempaa keskustelua tekijöistä, tavoitteista ja käyttötilanteista. Lisäksi eettistä dilemmaa on analysoitava yhdessä oleellisten sidosryhmien, suunnittelijoiden ja eettisten asiantuntijoiden kanssa. Tämä auttaa ymmärtämään, mitkä eettiset periaatteet ja arvot määrittävät käsiteltävän tekoälysovelluksen rajat.

14.2 Tekoäly muutosvoimana

14.2.1 Tekoäly haastaa luottamusta

Kun julkishallinnossa kehitetään tekoälyratkaisuja, otetaan samalla vahvasti kantaa siihen, millaista tulevaisuuden yhteiskuntaa halutaan rakentaa. Mikään muu teknologia ei ole ihmiskunnan historiassa vastaavalla tavalla horjuttanut kansalaisten luottamusta samalla, kun se on avannut huikeita kehitysnäkymiä ihmisen hyvinvoinnin parantamisesta. Keskustelut heijastavat kasvavaa huolta ihmisen hallinnan menettämisestä tekoälykehityksessä: siitä, millä tavoin päätöksentekoalgoritmit voivat heikentää julkisen päätöksenteon legitiimiyttä ja yhteiskunnallista solidaarisuutta, lisätä maailmanlaajuista vallan ja vaurauden epätasapainoa ja syventää eriarvoisuutta.

Tekoälyjärjestelmien kehittämistä voidaankin pitää poliittisena, valtaa ja legitiimiyttä koskevana ilmiönä. Kyse on sekä kansalaisten ja hallitusten välisestä, että maallikoiden ja asiantuntijoiden välisestä luottamuksesta oleellisena yhteiskunnallisena haasteena. Luottamus on se kulmakivi, jonka varaan tekoälysovellusten kehittäminen ja käyttö yhteiskunnassa tulisi rakentaa. **Se rakentuu kuulemisesta ja osallisuudesta, mukaan ottamisesta, vaikuttamisesta ja demokratiasta. Ilman luottamusta julkiseen päätöksentekoon, tahtoon ja harkintaan ei tekoälyn potentiaalia voida täysin hyödyntää.** Tekoäly haastaa siis paitsi ihmiskunnan hyvän toteutumisen ja ihmisten kukoistamisen myös **tavat, joiden avulla sen kehitystä koordinoidaan ja hallinnoidaan.** Siksi on hyvä kysyä, miten nykyisiä hallintorakenteita ja -välineitä voisi kehittää edistämään luottamuksen rakentumista sekä tukemaan sitä, että tekoälyn haasteista ja ratkaisuista syntyy yhteiskunnassa yhteinen ymmärrys.

14.2.2 Lisää näkemystä kokonaisuuden hallintaan

Kuten edellä on mainittu, tekoälyn hyödyntämisessä painottuvat sosiaalisiin kontekseihin ja ihmisten käyttäytymiseen liittyvät ongelmat. Euroopan komission asettama High-level Expert Group on Artificial Intelligence (AI HLEG) toteaa, että koska tekoäly voi parhaimmillaan olla ihmiskunnan kukoistusta tukeva teknologia, tulisi ymmärtää juuri yhteiskunnallisesta näkökulmasta, miten tekoälyä voidaan hyödyntää, jotta sillä edistetään ihmisen hyvää ja parannetaan ihmisten elämänlaatua. Siksi erityisesti julkishallinnon alalla tarvitaan tutkimusta, joka tunnistaa tekoälykehityksen **systeemisen luonteen** ja tuo **vaikutusten ennakoivan arvioinnin** mukaan kehitykseen.

Tekoälyn käytön vaikutuksia yhteiskunnassa, julkisessa hallinnossa ja etenkin kunnissa on toistaiseksi selvitetty vain vähän osittain siitä syystä, että tekoälyn hyödyntäminen on pääosin vasta suunnittelun tasolla eikä tekoälyä ole vielä otettu käyttöön laajamittaisesti. Osittain tästä syystä nykyinen tekoälyyn perustuvien toimintojen suunnittelu ja ohjaus on monilta osin epäjohdonmukaista ja hajanaista. Tekoälyn hyödyntämisen esteenä ovat muun muassa **liiallinen varovaisuus, osaamisen puute ja epätietoisuus tavoista hyödyntää** tekoälyä eettisesti ja vastuullisesti. Tekoälyteknologioiden avulla pyritään ratkaisemaan monimutkaisia ongelmia, mutta haluttujen vaikutusten aikaansaamista vaikeuttaa se, ettei kokonaisuutta hallita ja toiminnan reunaehdoista vallitsee epätietoisuus. Nämä puutteet hidastavat uusien teknologioiden käyttöönottoa ja hyödyntämistä sekä julkishallinnossa että kaupallisissa innovaatioissa.

Kuntasektorilla koetaan tekoälyn käyttöönottoa hidastaviksi tekijöiksi sääntelyn, ymmärryksen, osaamisen ja koulutuksen puute sekä datan puutteellinen laatu. Myös vähäiset resurssit kunnissa sekä tekoälykeskustelun nostamat eettiset huolenaiheet tunnistetaan käyttöönottoa hidastaviksi. Kunnissa ei juurikaan ole järjestetty koulutusta tekoälyjärjestelmien hyödyntämisestä, ja osaamistarpeen ja resurssien suuntaamisen välillä koetaan ristiriita. Edelläkävijäorganisaatioiksi mielletään ne harvat, jotka ovat panostaneet koulutukseen.[4]

14.2.3 Lainsäädännön rinnalle pehmeää ohjausta

Lainsäädännön ja niin sanotun kovan sääntelyn rinnalle tarvitaan erilaisia käytännön ajattelutyökaluja ja käsitteellisiä kehikkoja, jotka auttavat yrityksiä ja julkisia toimijoita hahmottamaan moniulotteisia vastuullisuuteen ja etiikkaan liittyviä kysymyksiä. Positiivisen eettisen vaikutuksen maksimoimiseksi tarvitaan myös **eettistä johtamista sekä innovatiivisia työkaluja ja kehitysalustoja**, jotta tulevaisuuden tarpeet ja tekoälyn vaikutukset ymmärrettäisiin.

Vastuullinen tekoälyn kehittäminen edellyttää myös kansalaisyhteiskunnan integroimista aidosti tekoälykehitykseen. Niin viime aikoina julkaistu tekoälykirjallisuus kuin esimerkiksi OECD patistavat panostamaan avoimuuteen ja sidosryhmien osallistamiseen tekoälykeskustelussa. OECD katsoo, että kun sidosryhmät osallistetaan päätöksentekoon, se lisää hallinnon vastuuvelvollisuutta, laajentaa kansalaisten vaikutusmahdollisuuksia ja vaikutusvaltaa päätöksenteossa, kehittää kansalaisvalmiuksia, parantaa päätöksenteon tietopohjaa, vähentää täytäntöönpanokustannuksia ja saa toimijat hyödyntämään laajempia verkostoja innovaatiotoiminnassa ja poliittisessa päätöksenteossa.[5]

OECD:n luottamusarvioinnin mukaan[6] Suomessa on korkea luottamus julkiseen hallintoon mutta kokemus vaikutusmahdollisuuksista ja päätöksenteon ymmärtämisestä on suhteellisen matala. Avoimuutta sekä osallistavampaa valmistelua ja päätöksentekoa korostava toimintatapa edellyttää oikeanlaista organisaatioiden eettistä agendaa, sen

4 ETAIROS-tutkimushanke.
5 OECD (2020). Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave.
6 OECD (2021). Drivers of Trust in Public Institutions in Finland.

varmistamista, että tietämys ja eettinen visio tekoälystä ovat oikealla tasolla. Julkiselta toimijalta edellytetään ennen kaikkea uskottavuutta sekä hyvin perusteltuja, realistisia ja toteuttamiskelpoisia näkemyksiä.

14.2.4 Suunnittelijan vastuu on suuri

Tekoälyn eettinen suunnittelu ja suunnitteluosaaminen voidaan helposti nähdä erilaisten ”teknisten” osa-alueiden huomioimisena. Tällöin keskitytään arvioimaan lähinnä järjestelmän läpinäkyvyyttä (engl. *transparency*) ja vastuuvelvollisuuden toteutumista (*accountability*). Kun kehitetään eettisesti kestävää tekoäly-yhteiskuntaa, tärkeää on tietoisuuden herättäminen kansalaisten keskuudessa. Kykyä tunnistaa ne avaintekijät, jotka vaikuttavat algoritmin päätöksentekoprosessiin, kutsutaan selitettävyydeksi (*explainability*). Oikeudellisessa keskustelussa korostuu näiden osa-alueiden huomioiminen nostamalla esiin teknisten menetelmien, teollisuusstandardien, sertifikaattien, auditointimenetelmien ja vaikutusten arvioinnin merkitys. On kuitenkin tärkeää huomata, että kysymystä algoritmisen päätöksenteon toivottavuudesta yhteiskunnassa ei tulisi käsitellä yksin vastuuvelvollisuuden kautta. **Kysymys algoritmisten päätöksentekojärjestelmien laajemmasta vastuullisuudesta (*responsibility*) on yhä tärkeämpi tekoälyn etiikassa**.

Vastuullinen tarkastelukulma perustuu toisiaan täydentäviin näkökohtiin:

- **ennakointiin**, jossa mietitään kehitys- ja innovaatiotoiminnan sekä tavoiteltuja että tahattomia seurauksia
- **oman tekemisen reflektointiin,** jossa pohditaan niitä kehitystyön taustalla olevia olettamuksia ja sitoumuksia, jotka voisivat vaikuttaa kehitystyöhön
- **sitouttamiseen**, jossa otetaan relevantit sidosryhmät mukaan avoimeen vuoropuheluun tekoälykehityksen halutuista ja ei-halutuista seurauksista ja taataan haavoittuvien ja heikoimmassa asemassa olevien ihmisten tasa-arvoinen osallistuminen
- **vastaanottavuuteen ja reagointiin**, jossa muokataan työtä oikeaan suuntaan saadun palautteen, näkemysten ja arvojen mukaisesti.

Näiden prosessien avulla on mahdollista ymmärtää, miten suunnittelija tai kehitystiimi pystyy yhtäältä ennakoimaan eettisiä ristiriitoja ja toisaalta tarvittaessa muuttamaan työnsä etenemistä kehitysprojektin aikana.[7]

14.3 Viestit päättäjille ja kansalaisille

Tarvitaan yhteinen koordinoitu käsitys vastuullisesta tekoälystä. Koska tekoälyn soveltaminen ei ole ainoastaan niin sanotusti tekninen kysymys vaan laajemmin yhteiskunnallinen oikeudenmukaisuus- ja valtakysymys, se nostaa esille keskeisenä yhteiskunnallisena haasteena kansalaisten luottamuksen julkiseen hallintoon ja maallikoiden luottamuksen asiantuntijoihin. Hallinnonaloilla on oltava yhteinen ja koordinoitu käsitys

7 FCAI Ethical Advisory Board. Ethics Exercise Tool

tekoälyn toteuttamisesta ja sääntelystä. Tämä edellyttää koordinointia eri hallintotasojen välillä sekä sosiaalisia ja teknisiä rakenteita, jotka takaavat, että toiminta on avointa, ennustettavaa ja korjattavaa ja että eri näkökulmat otetaan huomioon. Tekoälyyn kohdistuvaa luottamusta voidaan edistää myös kuntien itsesääntelyn avulla.

Julkisen, yksityisen ja kolmannen sektorin yhteistyötä pitää tehostaa. Jotta tekoälyjärjestelmiä voidaan kehittää ja soveltaa, on lisättävä tietoisuutta tekoälykehityksen systeemisestä luonteesta sekä toimijoiden ja yhteiskuntien keskinäisistä riippuvuuksista ja jännitteistä maailmanlaajuisesti. Vastuuta reiluudesta ja oikeudenmukaisuudesta ei voida asettaa vain yksilöiden tai yksittäisten organisaatioiden hartioille. Tekoälyn eettinen käyttö vaatii kollektiivista yhteistyötä ja jatkuvaa keskustelua. Käytännössä tämä voi tarkoittaa esimerkiksi tarvetta luoda uudenlaisia sosiaalisia sopimuksia, jotka huomioivat algoritmisen päätöksenteon vaikutukset tulevaisuudessa.

Tarvitaan kansalaisten osallisuutta. Jotta voidaan korostaa kansalaisten roolia aktiivisina kehittäjinä, on lisättävä viestintää, vastavuoroisuutta, osallisuutta ja avoimuutta yhteiskunnan, kansalaisten ja erilaisten data- ja palveluekosysteemien välillä. Julkishallinnon tehtävänä on herättää, mahdollistaa ja ylläpitää kansalaiskeskustelua. Tämä voi toteutua esimerkiksi järjestämällä digitaalisia kansalaisalustoja. Näillä alustoilla kansalaisilla on mahdollisuus ymmärtää ja analysoida, mitä algoritmit todella tekevät ja millaisia vaikutuksia sillä on yhteiskuntaan. Sidosryhmäpaneelien avulla voidaan osallistaa keskusteluun esimerkiksi sukupuoliltaan, kulttuuritaustoiltaan tai osaamiseltaan monimuotoisia ryhmiä. Kansainväliset kokemukset näistä technology assessment -paneeleista ovat jo pitkään olleet positiivisia.

Tarvitaan lisää empiiristä ja selittävää tutkimusta. Tekoälyn käyttötapausten systemaattinen tarkastelu eettisestä näkökulmasta on silmiä avaavaa ja tuo eteemme eettisiä ulottuvuuksia, joita muuten olisi mahdotonta hahmottaa. Tätä kontekstuaalista ymmärrystä tukemaan tarvitaan lisää empiiristä ja selittävää tutkimusta tekoälyn käytöstä, läpinäkyvyyttä tutkimustietojen jakamisessa sekä tietoa parhaista käytännöistä riskien hallitsemiseksi. Erityisesti hallintoalalla tarvitaan tutkimusta, joka tunnistaa eri yhteisöjen kulttuuriset herkkyydet ja tuo nämä yhteisöt mukaan tekoälykehitykseen.

On panostettava koulutukseen. Suunnittelun ja suunnitteluosaamisen kehittäminen on oleellista. Lisäksi on opetettava ihmiset ymmärtämään tekoälyn etiikkaa kaikissa väestöryhmissä. Tekoälyn eettinen ymmärrys pitäisi nähdä monialaiseen koulutukseen perustuvana yleisenä kansalaistaitona, joka kannustaa ja antaa valmiuden itsesääntéleviin toimiin.

Meidän on muutettava fokusta. Sorrumme tällä hetkellä helposti etsimään tekoälyn eettisiä ongelmia. Sen sijaan pitäisi kysyä, miten tekoälyä voidaan maksimaalisesti hyödyntää, jotta sillä edistetään ihmisten hyvää ja parannetaan heidän elämänlaatuaan. Tekoälyratkaisujen ennakointia pitäisi siksi parantaa. Samalla olisi **opittava arvioimaan teknologian hyödyntämistä yhteiskunnallisesti, moninäkökulmaisesti ja arvojen näkökulmasta**. Tässäkin olennaista on osallistaa sidosryhmiä keskusteluun tekoälyn toivotuista vaikutuksista mahdollisimman laajasti, jotta voidaan saada aikaan merkityksellinen lopputulos.

14.4 Tekoäly-yhteiskunnan etiikkaa

Yhä useammat ratkaisut kansalaisten arkea koskevassa julkisessa päätöksenteossa perustuvat tekoälyn toteuttamaan analyysiin. Tämän odotetaan näkyvän parempana arjen sujuvuutena, mutta onko kolikolla myös toinen puoli? Ohjaako tekoäly käyttäytymistämme ja jopa elämämme suuntaa? Tiedetään, että tekoälyn hyödyntämisellä on suoria ja epäsuoria vaikutuksia kansalaisten kohteluun, oikeuksiin ja mahdollisuuksiin. Maailmanlaajuisessa tarkastelussa huolta ovat herättäneet muun muassa datan keruun monopolit, uhat yksityisyyden loukkaamiselle ja datan väärinkäytölle, kansalaisten valvonta, sosiaalisen median kuplat ja sen käytön vaikutukset, verkkopalveluiden mahdollistama laiton toiminta sekä digitalisoitumisen vaikutukset energiankulutukseen. On kannettu huolta jopa siitä, miten tekoäly vaikuttaa ihmisten maailmankuvaan.

Tekoälyjärjestelmien suunnittelu ja käyttö eivät siis ole pelkästään moniteknologista toimintaa. Ne sisältävät sosiaalisia, psykologisia, taloudellisia, poliittisia ja oikeudellisia näkökohtia, joilla on syvällisiä vaikutuksia yksilöihin, yhteiskuntaan ja ihmiskuntaan. Siirtymä tekoäly-yhteiskuntaan ja tämän prosessin ohjaaminen ovatkin aikamme suurimpia haasteita, joissa korostuvat kysymykset tekoälyjärjestelmien systeemisestä luonteesta sekä vastuullisesta ja eettisestä käytöstä. Yritykset sisällyttää etiikka tekoälyjärjestelmien kehitystyöhön ovat toistaiseksi olleet melko pinnallisia, sillä eettiset periaatteet ja kehykset ovat suhteellisen väljästi tulkittavia ja siksi vaikeita soveltaa käytäntöön. Eettinen päätöksenteko on lisäksi kontekstista riippuvaista, mikä haastaa etiikan perinteistä, periaatelähtöistä suunnitteluajattelua.

On kuitenkin olemassa yhteinen näkemys, että tekoälyn tulisi edistää hyvää yhteiskunnassa ja että sen käyttöä pitäisi koordinoida oikeudenmukaisesti, puolueettomasti ja avoimesti. Miten tämä toteutetaan? Julkinen keskustelu ja tutkimus sosioteknisestä muutoksesta on lisääntynyt, mutta tekoälyn eettinen potentiaali on vasta hahmottumassa.

15 Tekoälystä tietoturvallista

Antti Honkela
apulaisprofessori, TkT
Tietojenkäsittelytieteen osasto
Helsingin yliopisto

15.1 Mistä on kyse?

Yleisesti käytössä olevat tietokoneet ja muut tietotekniset laitteet ovat pohjimmiltaan koneita, jotka toistavat niille täsmällisesti ohjelmoituja tietojenkäsittelyoperaatioita erittäin nopeasti. Tekoälyjärjestelmät hyödyntävät näiden laitteiden nopeutta uudenkaltaisissa sovelluksissa, joissa järjestelmän toimintaa ei ole ohjelmoitu yksityiskohtaisesti suoraan vaan esimerkiksi aiemmista esimerkeistä yleistämällä tai monimutkaisen optimointiongelman ratkaisuna. Aiemmista esimerkeistä yleistämiseen perustuva koneoppiminen on mahdollistanut viimeisten kymmenen vuoden aikana valtavia edistysaskelia esimerkiksi konenäössä, puheentunnistuksessa ja luonnollisen kielen ymmärtämisessä ja tuottamisessa tietokoneella.

Keskityn kirjoituksessani kolmeen koneoppimiselle ominaiseen turvallisuuskysymykseen: opetusaineiston tietosuojaan ja yksityisyyden suojaan, tekoälyjärjestelmien tietoturvaan sekä järjestelmien turvallisuuteen vuorovaikutuksessa ihmisten kanssa. Samoja riskejä käsitellään tarkemmin Liikenne- ja viestintävirasto Traficomin tuoreessa selvityksessä.[1]

Koneoppimista käyttävien tekoälyjärjestelmien kehittämisessä tarvitaan suuria opetusaineistoja. Monet kiinnostavat tekoälyn sovellukset esimerkiksi terveydenhuollossa liittyvät ihmisiin, jolloin tarvittavat opetusaineistot sisältävät henkilötietoja. Näiden aineistojen keräämiseen ja käsittelyyn sekä niistä johdettujen järjestelmien käyttöön liittyy tietosuojariskejä, ja ne saattavat loukata aineiston kohteiden yksityisyyttä.

Tekoälyjärjestelmät ovat tietojärjestelmiä. Kuten kaikissa tietojärjestelmissä, niissä voi olla tietoturvaheikkouksia. Tämä pitää huomioida, jos niitä käytetään turvattomassa ympäristössä, kuten internetissä, tai jos epäluotettavilla tahoilla on mahdollisuus manipuloida niiden kehittämisessä käytettyjä opetusaineistoja tai muita syötteitä.

Tekoälyjärjestelmien ”älykkyys” poikkeaa merkittävästi ihmisen älykkyydestä. Ne voivat tehdä systemaattisesti ihmisen näkökulmasta järjettömiltä vaikuttavia virheitä. Sovelluksissa, joissa virheistä voi olla merkittävää haittaa tai vaaraa, onkin tärkeää rakentaa mekanismeja näiden riskien ehkäisemiseksi hyödyntämällä esimerkiksi ihmisen valvontaa.

1 A. Vähä-Sipilä, S. Marchal & M. Aksela. *Tekoälyn soveltamisen kyberturvallisuus ja riskienhallinta*. Traficomin tutkimuksia ja selvityksiä 9/2021.

15.1.1 Tekoäly ja tietosuoja sekä yksityisyyden suoja

Tyypillinen oppiva tekoälyjärjestelmä kuvaa syötteitä vasteille, esimerkiksi kuvan tiedoksi siinä näkyvistä liikennemerkeistä tai henkilön terveystiedot arvioksi nykyisestä tai ennusteeksi tulevasta terveydentilasta.

Tällaisten järjestelmien rakentamiseen tarvitaan suuria esimerkkiaineistoja esimerkiksi suuren ihmisjoukon terveystiedoista ja havaitusta terveydentilasta, joista järjestelmä oppii näiden välisen yhteyden. Monissa kiinnostavissa sovelluksissa tarvittavat aineistot sisältävät henkilötietoja, kuten arkaluontoisia terveystietoja. Näiden käsittelyyn liittyy useita tietosuojariskejä, joita ei voida poistaa helposti esimerkiksi poistamalla tiedoista suorat tunnisteet:

1. Jos aineisto kerätään yhteen tietokantaan, riski tietojen paljastumiseksi kasvaa, mikäli tietokannan suojaus pettää.
2. Aineiston osia voi paljastua järjestelmän opetuksen aikana opetukseen osallistuville ulkopuolisille tahoille.
3. Osia aineistosta voi olla mahdollista selvittää valmiiksi opetetusta järjestelmästä.

Kuvattuja riskejä voidaan ehkäistä erilaisin teknisin ratkaisuin. Järjestelmiä on mahdollista opettaa käyttämällä hajautettua dataa, jolloin vältetään keskitettyyn tietokantaan liittyviä riskejä. Esimerkiksi Google ja Apple ovat opettaneet mobiilikäyttöjärjestelmiinsä käyttäjien kirjoituksia hyödyntäviä malleja, joiden opetuksessa käytetty data on pysynyt käyttäjien laitteilla eikä ole paljastunut edes mallin kehittäjille. Samaa periaatetta voidaan hyödyntää esimerkiksi silloin, kun yhdistetään terveystietoja eri maista, jos esimerkiksi lainsäädäntö ei salli datan luovutusta. **”Algoritmi tulee datan luo”** olisikin usein hyvä periaate parantamaan tietosuojaa.

Aineiston paljastumista opetuksen aikana voidaan estää salausmenetelmiä käyttämällä. Sopivilla menetelmillä on mahdollista opettaa järjestelmä salattua dataa käyttämällä siten, että kehittäjä ei voi nähdä dataa. Hintana näillä ratkaisuilla on usein se, että järjestelmän opettamiseen tarvittava laskenta-aika on merkittävästi suurempi.

paljastuminen valmiista järjestelmästä on hankalampi ongelma, johon perinteiset salausmenetelmät eivät pure. Eduskunnan tulevaisuusvaliokunnassa esiintyi hiljattain GPT-3-malliin pohjautuva tekstiä tuottava tekoälyjärjestelmä. Tutkijat ovat osoittaneet, että GPT-3:n edeltäjä GPT-2 on mahdollista houkutella toistamaan täydellinen kopio opetusaineistossa esiintyneistä henkilötiedoista[2]. Sama ilmiö on havaittu muillakin testatuilla järjestelmillä. Ilmiötä voidaan estää hyödyntämällä differentiaalista tietosuojaa, mutta hintana on usein se, että mallin tarkkuus heikkenee. Ilmiön tarkempaan ymmärtämiseen tarvitaan lisätutkimusta. **Henkilötietoja tai muita luottamuksellisia tietoja käyttämällä opetettujen tekoälyjärjestelmien jakamisessa on kuitenkin syytä olla varovainen.**

---

2 N. Carlini et al. *Extracting Training Data from Large Language Models*. In Proceedings of the 30th USENIX Security Symposium, 2021.

15.1.2 Tekoälyn tietoturva

Tekoälyjärjestelmät ovat tietojärjestelmiä ja siten alttiita samoille tietoturvariskeille kuin muutkin tietojärjestelmät. Erityisesti opetettaviin järjestelmiin liittyy lisäksi niille ominaisia riskejä, joista on seuraavassa muutama esimerkki:

1. Opetusaineistoa manipuloimalla järjestelmän kehittäjä tai ulkopuolinen hyökkääjä voi upottaa järjestelmään takaoven, jota on käytännössä mahdoton havaita mutta joka saa järjestelmän toimimaan valituissa tilanteissa hyökkääjän haluamalla tavalla. Esimerkiksi hakemusten käsittelyyn käytetty järjestelmä voitaisiin manipuloida hyväksymään tai hylkäämään aina hakemukset, joissa on tietty harmittomalta vaikuttava avainsana tai muu piirre.
2. Valmiista tekoälyjärjestelmästä on mahdollista löytää heikkouksia, joissa jopa ihmiselle huomaamattomat muutokset saavat järjestelmän toimimaan täysin eri tavalla. Tutkijat ovat kehittäneet esimerkiksi liikennemerkkeihin kiinnitettäviä tarroja, jotka eivät erityisesti häiritse ihmistä mutta joiden takia liikennemerkkien tunnistukseen käytetyt järjestelmät alkavat luulla stop-merkkiä nopeusrajoituksen merkiksi[3]. Muiltakin aloilta on vastaavia esimerkkejä siitä, miten on mahdollista tietoisesti kehittää tapoja, jotka eivät hämäisi ihmistä mutta joilla tekoälyä on mahdollista huijata.
3. Mikäli järjestelmä oppii adaptiivisesti toiminnan aikana, käyttäjät voivat ohjata sen toimintaa epätoivottuun suuntaan. Esimerkiksi sosiaalisen median keskusteluista oppinut Microsoftin Twitter-botti muuttui alle vuorokaudessa ystävällisestä aggressiiviseksi rasistiksi muiden käyttäjien manipuloinnin vuoksi[4].

Yllä kuvattuja riskejä voidaan ehkäistä yhdistelmällä teknisiä ja järjestelmätason ratkaisuja. **Tehokkain tapa ehkäistä opetusaineiston manipulointia on käytetyn aineiston tarkka laadunvalvonta. Valmiin järjestelmän käytössä tämä ei ole yhtä helppoa, jolloin on tärkeää lisäksi vahvistaa järjestelmää niin, että se ei olisi niin altis huijausyrityksille** (vrt. liikennemerkkeihin liimattavat tarrat edellä). Järjestelmien vahvistamiseen soveltuvat teknologiat ovat tällä hetkellä hyvin aktiivisen tutkimuksen kohteena, eikä ole vielä selvää, mikä on lopulta mahdollista ja mikä ei. Kuten monessa turvallisuuskysymyksessä, tässäkin on riskinä hyökkääjien ja puolustajien välinen varustelukilpailu, jossa kehittyneempää puolustusjärjestelmää seuraa sen kiertävä uusi, vielä edistyneempi hyökkäys.

Useimmat koneoppimiseen pohjautuvat tekoälyjärjestelmät eivät opi enää käytön aikana. Tämä toisaalta helpottaa niiden toiminnan ymmärtämistä, toisaalta rajoittaa mahdollisuuksia sopeutua muuttuvaan ympäristöön. Käytön aikana tapahtuvan oppimisen

---

3 K. Eykholt et al. *Robust physical-world attacks on deep learning visual classification*. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition CVPR 2018). IEEE, 2018.

4 O Schwartz. "*In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation*". IEEE Spectrum 25.11.2019. Viitattu 26.11.2021.

turvallisuus on toistaiseksi hyvin nuori tutkimusala. Tehokkaita ratkaisuja tämän alueen ongelmiin saadaan todennäköisesti odottaa vielä useita vuosia.

15.1.3 Ihminen tekoälyn valvojana ja yhteistyökumppanina

Tekoälyjärjestelmiä käytetään yhä enemmän sovelluksissa, joilla on merkittäviä vaikutuksia esimerkiksi ihmisten talouteen, hyvinvointiin ja terveyteen. Turvallinen tekoäly vaatii mekanismeja, joilla varmistetaan, että nämä vaikutukset eivät ole haitallisia. Nämä kysymykset ovat läheisessä yhteydessä tekoälyn vastuullisuuteen ja etiikkaan, mutta osa kysymyksistä liittyy myös turvallisuuteen: miten hallitaan autonomisten ajoneuvojen liikenneturvallisuusriskejä tai varmistetaan, ettei tekoälyn ehdottama hoitotoimenpide aiheuta tarpeetonta vaaraa?

Tekoälyjärjestelmille ja ihmisille on yhteistä se, että molemmat tekevät virheitä ja erehdyksiä. Tekoälyn ja ihmisten tekemät virheet ovat kuitenkin usein erilaisia, jolloin niitä voidaan ehkäistä koneen ja ihmisen yhteistyöllä. Matkalla on monta haastetta:

- Miten selitetään koneen toimintaa ihmiselle niin, että tämä voi sitä arvioida ja havaita mahdollisia ongelmia?
- Miten kehittää järjestelmiä, jotka ovat ihmiselle helpommin ymmärrettäviä, menettämättä niiden tehokkuutta?
- Miten rakennetaan ihmiselle mahdollisuus valvoa koneen toimintaa niin, että tehtävä pysyy mielekkäänä ja ihminen pystyy oikeasti puuttumaan toimintaan tarpeen vaatiessa riittävän nopeasti?
- Miten järjestetään ihmisohjaus menettämättä liikaa koneen tehokkuushyötyjä?
- Miten koulutetaan ihmisiä seuraamaan koneiden toimintaa tehokkaasti?

Kuten tavalliset tietojärjestelmät, tekoälyjärjestelmät eivät opi virheistään, ellei niitä ole erityisesti siihen suunniteltu. Yleisin tapa korjata virheitä on uuden version julkaisu ja käyttöönotto. Tämä rajoittaa mahdollisuuksia nopeaan reagointiin, erityisesti jos käyttäjä ja kehittäjä ovat erillisiä tahoja. Aiemmin mainittu aktiivinen oppiminen käytön aikana mahdollistaisi teoriassa nopeamman reagoinnin, mutta näiden menetelmien turvallisuudessa on vielä paljon avoimia ongelmia.

15.2 Tekoäly muutosvoimana

Erilaiset tekoälyjärjestelmät ovat viime vuosina tulleet yhä enemmän osaksi ihmisten arkea erilaisissa nettipalveluissa ja elektronisissa laitteissa. Niiden yleistymisen voi olettaa jatkuvan ja laajenevan uusille aloille. Paljon potentiaalia on esimerkiksi terveydenhuollossa. Toimiessaan nämä järjestelmät sujuvoittavat arkea ja tehostavat palveluita, mutta samalla ne luovat uusia turvallisuusuhkia.

Tekoälyjärjestelmät ovat ensisijaisesti tietojärjestelmiä, ja näihin molempiin pätevät monet samat lainalaisuudet. Perinteisten tietojärjestelmien turvallisuudesta muodostuu valitettavan usein ”kisa pohjalle”, jossa järjestelmien rakentajat kilpailevat kustannuksia leikatakseen säästämällä turvallisuudesta, koska ostajat eivät osaa sitä vaatia eivätkä ole

valmiita siitä maksamaan. **Yksi osa ongelmaa on se, että turvallisuuden puutteet koituvat usein muiden vahingoksi. Kuinka moni auton ostaja olisi vapaaehtoisesti valmis maksamaan enemmän autosta, joka olisi kolaritilanteessa turvallisempi muille, esimerkiksi jalankulkijoille?**

Turvallisuuteen liittyvissä sovelluksissa hyökkääjän ja puolustajan välille muodostuu helposti varustelukierre, jossa paremman puolustuksen hyöty häviää paremman hyökkäyksen edessä, mikä edellyttää vielä parempaa puolustusta ja niin edelleen. Tällaisia kierteitä voi syntyä myös silloin, kun tekoälyn antamaa turvaa murretaan ja puolustetaan. Tekoäly on vahva työkalu kisan molemmilla puolilla, jolloin molempien osapuolten voi olla käytännössä pakko hyödyntää sitä pärjätäkseen kilpailussa. Tieteellinen perustutkimus voi joissain tilanteissa tarjota ulospääsyn varustelukierteestä. **Esimerkiksi differentiaalista tietosuojaa käyttämällä voidaan todistaa matemaattisesti, että yksikään hyökkääjä ei voi loukata käyttäjien tietosuojaa, vaikka tämä käyttäisi mitä tahansa tulevaisuuden teknologiaa.** Vastaavasti tekoälyjärjestelmien huijaamista tai kiertämistä vastaan kehitetään teknologioita, joilla voidaan todistaa mallin kestävän varmasti tiettyjä manipulaatioita.

Turvallisten tekoälyteknologioiden kehitys tuottaa todennäköisesti tulevaisuudessa helppokäyttöisiä työkaluja, joilla järjestelmien turvallisuutta voidaan vahvistaa ja varmistaa. **Turvallisen tekoälyn käytössä tullaan kuitenkin todennäköisesti törmäämään vaikeisiin arvovalintoihin, koska valinnoilla on hintansa.** Tietosuojaa kunnioittava tai manipulaatioita kestävä järjestelmä tuottaa usein optimiolosuhteissa vähemmän tarkkoja ennusteita, jolloin järjestelmän tuottama hyöty pienenee. **Mikä pahempaa, tämä tarkkuuden lasku vaikuttaisi kohdistuvan helposti voimakkaampana vähemmistöihin ja yleensä enemmistöstä poikkeaviin yksilöihin, mikä voi loukata yhdenvertaisuutta.** Näiden eri arvojen yhteensovittamisesta tarvitaan vilkasta yhteiskunnallista keskustelua, ettei ongelmien ratkaisu jää yksin teknologian kehittäjien vastuulle. Yhteensovittamisessa on paljonkin pelivaraa, koska turvallisuus ei ole päällä tai poissa -tyyppinen ominaisuus, vaan useimmissa teknologioissa on mahdollista säätää suojan tasoa ja siten myös kustannuksia ja sivuvaikutuksien voimakkuutta.

15.3 Viestit päättäjille ja kansalaisille

Tietoturvan ja -suojan sääntely on tietoyhteiskunnan ympäristönsuojelua. Koska kustannukset kohdistuvat ulkopuolisiin, markkinat eivät sitä ratkaise, vaan tarvitaan ulkopuolista ohjausta ja sääntelyä.

Turvallisen ja eettisen tekoälyn toivottavat ominaisuudet voivat olla ristiriitaisia, jolloin yhden toteuttaminen heikentää muita. Ominaisuudet eivät ole mustavalkoisesti päällä tai poissa, vaan niiden välistä tasapainoa voi yleensä säätää tarkemmin. Tarvitaan sopivasti vapautta ja sääntöjä, jotta löydetään kuhunkin sovellukseen paras ratkaisu. Samalla pitää varmistaa, että tasapainoa muodostettaessa huomioidaan myös muiden kuin järjestelmän kehittäjän ja omistajan etu.

Valmiista tekoälyjärjestelmästä on mahdoton todentaa, toimiiko se joka tilanteessa oikein eikä esimerkiksi sisällä piilotettua takaovea. Luotettavuuden varmistamiseksi on siksi tärkeää varmistaa prosessi, jolla järjestelmä on rakennettu.

Tekoälyjärjestelmät eivät ole erehtymättömiä, mutta eivät ole ihmisetkään. Näiden kahden tekemät virheet ovat luonteeltaan hyvin erilaisia. Tärkeissä sovelluksissa käytettäviä tekoälyjärjestelmiä tulee valvoa tavalla, joka on tarkoituksenmukainen kyseisessä sovelluksessa eikä turhaan nosta valvonnan kustannuksia mahdollisten virheiden haittoja suuremmiksi.

Datan luovuttamisen sijaan kannattaa suosia järjestelmiä, joissa koodi tulee valvotusti datan luo. Esimerkiksi sen sijaan, että terveystietoja luovutettaisiin ulkopuolisille, olisi parempi suosia ratkaisuja, joissa dataa ei luovuteta vaan sen tarvitsijat pääsevät käsittelemään tarvitsemiaan osia turvallisesti suojatussa ympäristössä.

Datan keskittämisen sijaan kannattaa mahdollisuuksien mukaan edistää datan hajauttamista tietosuojariskien pienentämiseksi. Suunniteltu kansallinen genomitietorekisteri on esimerkki erittäin suuririskisestä hankkeesta, jonka hyödyistä on erimielisyyttä.

Tekoälyn tietoturvan ja tietosuojan osaamista tulee ylläpitää ja kehittää aktiivisesti kaikilla tasoilla. Turvaa ei usein voi lisätä jälkikäteen valmiiseen järjestelmään, joten kehittäjien pitää osata rakentaa se järjestelmän sisään alusta asti, mikä vaatii laajaa koulutusta. Toisaalta tarvitaan uusien, turvallisempien menetelmien tutkimusta. Vain panostamalla alan perustutkimukseen Suomessa voidaan varmistaa paras osaaminen ja korkeatasoisin uusien menetelmien opetus.

15.4 Kansalaisten tietosuojan turvaava tietojen jako

Yhdysvalloissa järjestetään kymmenen vuoden välein väestönlaskenta (US Census). Laskennan tuloksena julkaistaan tilastoja eri ikä-, etnisiin ja muihin. ryhmiin kuuluvista henkilöistä tarkalla maantieteellisellä jaolla. Näitä tilastoja käytetään erilaisiin hallinnollisiin ja tutkimustarkoituksiin. Tietoja on aiemmin suojattu muun muassa korvaamalla pieniä ihmismääriä koskevia tietoja synteettisillä arvoilla. Väestönlaskentaviraston tutkimuksen mukaan näistä suojatoimista huolimatta huomattava osa väestöstä voitiin tunnistaa julkaistuista tiedoista yhdistämällä niitä muihin julkisesti saatavilla oleviin tietoihin.

Uusin vuoden 2020 US Census on ottanut tietojen suojaamisessa käyttöön differentiaalisen tietosuojan, joka perustuu tietojenkäsittelytieteen yksityisyyttä turvaavien teknologioiden tutkimukseen[5]. Differentiaalisen tietosuojan käyttö vähentää huomattavasti tunnistamisriskiä ja parantaa siten selvästi kansalaisten yksityisyyden suojaa. Sen käyttöä on silti myös kritisoitu, koska tietosuojan turvaaminen vaatii lisäämään tuloksiin enemmän satunnaisuutta, mihin tietojen käyttäjät eivät ole tottuneet ja mikä saattaa haitata joidenkin tietojen käytön vakiintuneiden päämäärien saavuttamista.

Differentiaalisen tietosuojan kaltaisten menetelmien käyttö avaa laajemminkin mahdollisuuksia yksityisyyden turvaavaan tietojen jakamiseen. Yhdysvalloissa saatavia kokemuksia on siis syytä seurata. Yksi erityisen kiinnostava mahdollisuus on differentiaalisen tietosuojan yhdistäminen niin sanottuun synteettiseen dataan, jossa esimerkiksi koneoppivan tekoälyn opettamiseen voidaan luoda todellista vastaava keinotekoinen aineisto, joka ei todistettavasti voi loukata kenenkään yksityisyyttä. Näillä menetelmillä on kuitenkin rajoitteensa, ja lisätutkimusta tarvitaan, jotta niiden potentiaali voidaan ymmärtää.

---

5 US Census Bureau. 2020 Census Data Products: Disclosure Avoidance Modernization. Viitattu 26.11.2021.

16 Tekoälystä fyysisesti turvallista: esimerkkinä automaattiset ajoneuvot

Matti Kutila
Tiimipäällikkö, Automaattiset ajoneuvot
VTT

16.1 Mistä on kyse?

Ajoneuvojen automaattisuus on kiehtonut tutkimustoimintaa aina vuodesta 1925, jolloin Houdina Radio Control Co. demonstroi New Yorkissa langattomasti ohjattua autoa. Euroopassa merkittävin lähtölaukaus oli autoteollisuuden EUREKA-projekti nimeltään Prometheus. Projektissa etsittiin sen aikaisella tietotekniikalla vastauksia automaattiajamisen haasteisiin ja tarvittaviin toiminnallisuuksiin[1]. Suuren yleisön tietoon ajoneuvojen automaatio tuli käytännössä 2000-luvulla USA:n DARPA-haastekisoissa, joiden voittajatiimin vuonna 2007 muodosti Carnegie Mellonin yliopisto. Voittajajoukkueen salaisuus oli

MOTIVAATIO AUTOMAATION EDISTÄMISEEN:

A. Autoteollisuus
1) turvalliset autot
2) kuljettajaa houkuttelevat uudet toiminnallisuudet ja palvelut
3) teknologiajohtajuus ja maine

B. Hyötyajoneuvot
1) tuottavuuden lisääminen automaatiolla ja erikoistilanteissa etäohjauksella (yksi kuljettaja voi valvoa montaa ajoneuvoa ja ajaa tarvittaessa yksittäisiä ajoneuvoja etänä)
2) monotonisten siirtoajojen automatisointi tehdasalueilla
3) työturvallisuuden parantaminen

C. Viranomaiset
1) liikenneturvallisuus eri liikennemuodoille (ml. kaksipyöräiset)
2) liikkumisen mahdollistaminen kaikille
3) joukkoliikenteen tehostaminen ja tasapuoliset liikkumismahdollisuudet

D. Auton ostajat
1) ajamisen helppous
2) ”yleinen trendikkyys”
3) liikenneturvallisuus
4) mahdollisuus käyttää ajamiseen kuluva aika työelämään tai sosiaalisten suhteiden edistämiseen

1 Wikipedia: History of self-driving cars. Viitattu 10.9.2021.

poikkeuksellinen neuroverkkoon pohjautuva ajoneuvon ohjausstrategia. Yliopistosta siirtyi sittemmin väkeä kehittämään paljon kuuluisuutta saanutta Googlen automaattiautoa, johon on tähän mennessä panostettu yli 1,1 miljardia dollaria.

Automaation kehitysaskeleet on jaettu standardin SAE J3016 mukaisesti eri tasoihin.[2]

0	**EI AUTOMAATIOTA** – Melkein kaikissa uusissa autoissa on varoitusviestejä, kuten törmäysvaroittimia.
1	**KULJETTAJAN TUKI** – Nykyiset edistyksellisemmät kuljettajan tukijärjestelmät, kuten kaistalla pysymisen tuki, kuuluvat tähän (ns. aktiiviset turvajärjestelmät).
2	**OSITTAINEN AUTOMAATIO** – Automaatio hoitaa tiettyjä tehtäviä kuljettajan rinnalla. Esimerkiksi ruuhka-avustin seuraa edellä olevaa ajoneuvoa ja huolehtii turvavälistä. Teslan Autopilot on tason 2 järjestelmä, mutta tietyissä olosuhteissa siinä on tason 4 toiminnallisuuksia.
3	**EHDOLLINEN AUTOMAATIO** – Kuljettajan pitää aina pystyä ottamaan auto hallintaansa, jos automaattijärjestelmä pettää. Tällaisiin ratkaisuihin on olemassa tekninen valmius, mutta käytännössä niitä ei ole markkinoilla epäselvien vastuukysymysten vuoksi.
4	**KORKEA AUTOMAATIO** – Automaatio hoitaa ajamisen, mutta kuljettaja saa puuttua automaation toimintaan. Automaation toiminta on rajattu tiettyihin liikennealueisiin (esim. moottoritie ja kampusalue). Tähän luokkaan voidaan laskea pikkubussit (esim. Navya, Easymile ja Auve Tech), jotka liikkuvat hitaasti rajatuilla alueilla.
5	**TÄYSI AUTOMAATIO** – Automaatio hoitaa täysin auton hallinnan, eikä kuljettajalla ole välttämättä edes mahdollisuutta puuttua siihen. Ajoneuvosta saattavat puuttua hallintalaitteet kokonaan. Tällaisia ratkaisuja ei ole markkinoilla, vaikka jotkin bussivalmistajat niin väittävät. Virallisessa muodossaan 5. tason automaatio on mahdottomuus, sillä ideaalinenkin maailma sisältää aina poikkeamia.

Operation Design Domain (ODD) on noussut jopa automaatiotasoja tärkeämmäksi muun muassa EU:n T&K-puiteohjelmassa.[3] ODD on toimintaympäristön määrittely johon automaatiojärjestelmä on suunniteltu. Siinä määritellään tarkemmin ne vaatimukset ja skenaariot, joissa automaattiajoneuvon on kyettävä toimimaan. ODD:n merkitys

---

2 Laakso, J., Vesanto, A. & Ritari, J. 2017. Selvitys automaattiajamisen edellyttämistä tiedoista ja kehittämistarpeista. Liikenne- ja viestintäministeriön julkaisuja 19/2017; SAE Standards News: J3016 automated-driving graphic update. Viitattu 10.9.2021.

3 Society of Automotive Engineers (SAE): The key to autonomous vehicle safety is ODD. Viitattu 10.9.2021.

korostuu erityisesti automaatiotasoilla 4 ja 5, koska luotettavan automaation toteuttaminen kaikkiin mahdollisiin olosuhteisiin ja tilanteisiin johtaa helposti siihen, että erilaisia toimintavaihtoehtoja on määriteltävä valtava määrä vaikka liikennesäännöt asettavat tietyt reunaehdot. Jokainen toimintamalli vaatii erilaisen ohjelmointi -algoritmin koodaamiseen ja siksi kaikissa olosuhteissa toimivan ajoneuvon rakentaminen on lähes mahdotonta vaan pitää valita ne ODD:t joihin ajoneuvo on tarkoitettu. Ongelmaksi muodostuvat erityisesti hyvin harvinaiset tilanteet. Tällainen harvinainen tilanne on esimerkiksi se, kun auto ajaa, luultavasti sumeilematta, tulvineelle tieosuudelle mutta jähmettyykin paikalleen, koska sen anturit peittyvät veteen. Toinen ongelma on ihmisen aikomusten ennakointi, mikä on tietokoneille hyvin vaikeaa. Tietokoneen on esimerkiksi vaikea hahmottaa, onko jalkakäytävän reunassa seisova ihminen aikeissa rynnätä tien yli vai odotteleeko hän rauhassa kyytiä. Pelkästään kaupungissa vasemmalle kääntyminen vaatii erilaisen päättelyalgoritmin kuin moottoritienopeudessa ohittaminen.

Tällä hetkellä markkinoille **on tulossa Teslan, Audin, Mercedeksen ja muiden vastaavien lisäksi monia tason 2 henkilöautoja**, jotka toimivat niin sanotuissa hyvissä keliolosuhteissa ja paikannuksen ollessa riittävän tarkka. Audi on tuomassa tason 3 ajoneuvoa markkinoille vuosina 2022–2023.

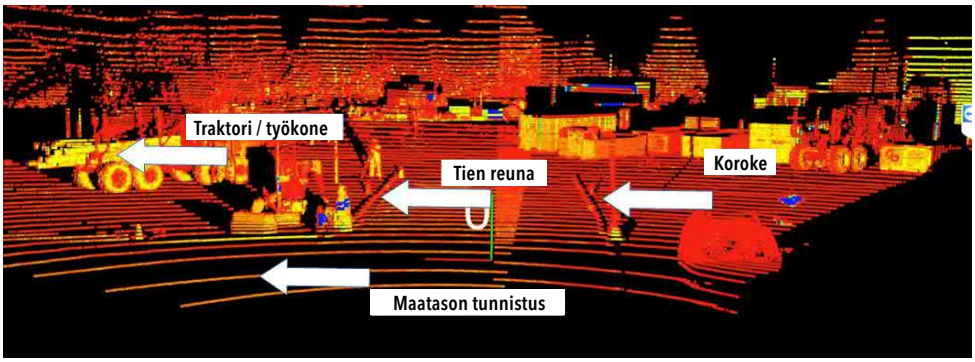
Saatavilla on **ensimmäisen/viimeisen kilometrin** (engl. *first/last-mile*)-tarkoitukseen sopivia tason 4 automaattiautoja, jotka kulkevat enintään 20 kilometriä tunnissa. Ne toimivat käytännössä aina ennalta opetetulla (ns. nauhoitetulla) reitillä eivätkä kykene itse luomaan uusia reititysvaihtoehtoja.

Useammalla valmistajalla on käytössä automaattisi**a työkoneita, joiden voi katsoa olevan tasolla 4**. Ne toimivat rajatuilla alueilla myös huonommissa olosuhteissa ja jopa ilman näkyviä kaistaviivoja. Nopeudet ovat pieniä. Nämä ajoneuvot toimivat erikoisluvilla muun muassa kävelyteillä, jolloin ne voivat poiketa yleisistä liikennesäännöistä.

Liikkuvissa työkoneissa ja erityisesti sisätilarobotiikassa on kehitetty ja sovellettu tekoälyä jo yli 30 vuotta. Alkuun tekoälyn käyttöä rajoittivat mikroprosessoreiden laskentakapasiteetti ja se, että niiden toimintaa ymmärrettiin heikosti. Käyttökohteita olivat muun muassa erilaisten kohteiden luokittelu ja laaduntarkistustehtävät teollisuudessa. Hyvin pian huomattiin, että niiden etu verrattuna klassiseen piirteisiin pohjautuvaan luokitteluun ei ollut iso. Ongelma neuroverkkoon pohjautuvassa tekoälyssä oli luokittelun käyttäytymisen hahmottaminen: matematiikka neuroverkon taustalla pohjautuu epälineaarisiin algoritmeihin. Niinpä käytännössä totuuksia vertailevat niin sanotut tikapuumallit (**if** pallo = sininen, **then** laita koriin 1, **elseif** pallo = oranssi, laita koriin 2 jne.) säilyivät ymmärrettävyytensä vuoksi valtavirtana aina vuoteen 2005 asti. Silloin hakukoneyhtiöt alkoivat käsittelemään massiivisia datamääriä tekoälyn avulla, mutta tikapuumalleja oli mahdotonta rakentaa niin suuren datamäärän hallintaan.

Autoteollisuuden lainsäädännöllinen pakote edellyttää, että autonvalmistajan pitää pystyä aukottomasti todistamaan, mitä auton ohjaustoiminnot tekevät saadessaan tietyn

syötteen. Hätäjarrujärjestelmä on yksi esimerkki: se saa aktivoitua vain, kun edessä on este eikä törmäystä voi välttää.. Neuroverkon kaltainen epälineaarinen musta laatikko ei ole paras tapa todistaa, miten auton ohjaustoiminnot reagoivat tietyissä tilanteissa. Se voi kuitenkin olla riittävä ratkaisu sellaisissa sovelluksissa, joiden turvakriittisyys on matala. Nyt tilanne on muuttunut, kun autoissa on useita korkearesoluutioisia datalähteitä tunnistamassa estettä (ks. kuva 1). Samoin kuin hakukoneyhtiöiden myös auton datalähteiden valtavan tietomäärän hallintaan on mahdotonta rakentaa tikapuumallia. Motivaatio tekoälyn hyödyntämiseen lähteekin nyt datan määrän aiheuttamasta pakosta eikä aiheen trendikkyydestä kuten 30 vuotta sitten.



KUVA 1. Hahmojen tunnistaminen laserskannerin tuottamasta 3D-kuvasta pohjautuu autoissa optisiin tai radiotaajuusantureihin ja niiden tuottamaa "kuvaa" tulkitsevaan tekoälyalgoritmiin.

16.2 Tekoäly muutosvoimana

Tekoäly avaa autoalalle uusia mahdollisuuksia mutta aiheuttaa myös haasteita:

- Tekoäly tuo **ajoneuvojen "alkeistason älykkyyden"** ja kyvyn sopeutua päättelyketjuja vaativiin tilanteisiin, joissa vastausta ei saada suoraan ohjelmoinnin perusfunktioilla (ns. *if–else-vertailull*a). Tämä on oikeastaan taistelu, joka on voitettava (engl. *must-win-battle*), jos halutaan ajoneuvon toimivan esimerkiksi Pohjois-Euroopan vaihtelevassa säässä.
- Tekoäly on ohjelmistotuote, joka on pakottanut **autoteollisuuden nopeuttamaan tutkimus- ja kehittämistoiminnan (t&k) sykliä ja avoimuutta** uusia toimijoita kohtaan uusien IT-alan toimijoiden (esim. Google, Baidu ja Uber) vanavedessä. Tämä uusi avoimuuden aika on iso mahdollisuus erityisesti kotimaiselle ohjelmistoteollisuudelle.
- **Älyajoneuvot poistavat rajoja kodin, työn ja liikkumisen välillä**. Ajoneuvo on paikka liikkua, työskennellä ja harrastaa sosiaalista elämää. Työskentelyä voidaan ajoneuvon liikkuessa jatkaa etätyöperiaatteella. Tämä mahdollistaa sen, että ihmiset voivat asua ja elää vaikkapa mökeillä ja harvaan asutuilla alueilla. Covid-19-pandemian myötä ihmisiä on siirtynyt kalliista Piilaaksosta halvemmille asuinalueille työskentelemään. Isot IT-alan trendit syntyvät usein Kaliforniassa, eikä ole syytä, miksi vastaava ilmiö ei leviäisi myös Suomeen.

- Tulevaisuuden automaattiautot maksanevat suhteessa palkkatasoon enemmän kuin nykyisin. Tämä on seurausta siitä, että autoja tehdään vähemmän ja tutkimus- ja kehittämistyötä enemmän. Massavalmistuksesta saatava komponenttien hintaetu vähenee. Sen vuoksi ajoneuvojen hinnat nousevat, vaikka sähköautoissa komponenttien määrä on pienempi kuin polttomoottoriautoissa. Tämän vuoksi **kaikki merkittävät ajoneuvovalmistajat panostavat nyt ajoneuvojen jako-, vuokraus- ja leasingtoimintaan, jossa ajoneuvoja vuokrataan omistamisen sijaan ja niillä ajetaan 24/7-periaatteella**. Autonvalmistajat alkavat keräämään tulovirtansa autojen käytöstä joka johtaa siirtymiseen isosta kertaostosta jatkuvaan kuluttajien menovirtaan. Siirtymä vähentää samalla radikaalisti parkkipaikkojen tarvetta, mutta ei toki poista ruuhkien aiheuttamia ongelmia.

Tekoäly tuottaa myös nipun uusia haasteita autoteollisuudelle:

- Autojen tyyppihyväksyntä ja todennettavuus vaikeutuvat. Mukautuva auto ei ole sama kuin ostohetkellä. Autojen turvallisen toiminnan todentamiseen tarvitaan erillisiä puolueettomia tarkistustahoja, joilla on vahva ymmärrys automaation toiminallisuudesta sekä tekniikkaan perustuvista syy-seuraussuhteista.
- Tekoälyalgoritmejä pitää kehittää ja validoida riittävän kattavasti siten, että eri sääolosuhteet, liikkumisrajoitteet ja etniset seikat (esim. ihmisten tunnistaminen ihonväristä ja pukeutumisesta riippumatta) tulevat huomioiduiksi. Mikä on riittävän kattava testiaineisto, ja paljonko tekoälyn virhetulkintoja voidaan sallia esimerkiksi 1 000:ta kilometriä kohti? Entä voidaanko ajoneuvo hyväksyä käyttöön maailmanlaajuisesti, jos se on testattu vain Kalifornian olosuhteissa?
- Autoja ei enää katsasteta vuosittain. Katsastaminen muuttuu online-toiminnaksi, jossa tulkitaan ajoneuvon diagnostiikkaa. Huolto- ja katsastustoiminnan raja käytännössä poistuu. Tulevaisuudessa ei tarvita niinkään puolueettomia katsastajia vaan puolueettomia huoltojen arvioijia.
- Kun ajoneuvot ovat verkottuneita ja itse ajattelevia, niiden väärinkäyttö muuttaa muotoaan. Nykyisen kaltaiset autovarkaudet loppuvat. Tulevaisuudessa ajoneuvojen väärinkäyttö edellyttää tietoteknisiä taitoja, mutta se mahdollistaa kyberhyökkäyksen jopa kokonaisiin ajoneuvojoukkoihin yksittäisen ajoneuvon sijaan Esimerkiksi Teslan kaikki autot ovat yhteydessä verkkoon ja niiden käyttömahdollisuuksia pystytään säätämään etänä. Tesla tai kolmas osapuoli voi säätää esimerkiksi auton toimintasädettä.
- Tekoäly monimutkaistaa ajoneuvoja. Sen seurauksena niihin saattaa tulla yllättäviä virhetoiminnallisuuksia, joita on käytännössä mahdotonta esitestata kattavasti – käyttötapausten joukkoon mahtuu miljoonia erilaisia harvoin tapahtuvia asioita. Hieman vastaavanlaisesta tapauksesta oli kyse, kun vuonna 1986 F-16-hävittäjästä löytyi ohjelmistovika. Vian vuoksi kone kääntyi ympäri aina, kun se lensi päiväntasaajan yli.[4]

---

4 Chicago Tribune 14.12.1986: *Software bugs turning deadly in complex era*. Viitattu 10.9.2021.

Osaltaan älyautojen ja tekoälyn ansiosta **Euroopan autoteollisuus yhdentyy**. Haasteena ei ole pitää yksittäisen maan, kuten automaa Saksan, vaan koko Euroopan kilpailukyvystä huolta erityisesti Kiinasta, Kanadan Ontariosta ja Yhdysvaltojen Kalifornia-ta tulevia kilpailijoita vastaan. Tähän eivät riitä Saksan tai Ranskan voimavarat, vaan se vaatii eurooppalaisia T&K-alliansseja, joihin niin yritykset kuin tutkimuslaitoksetkin osallistuvat. Tämä on jo nyt nähtävissä valtioiden sisällä. Vaikka Bosch, Mercedes, BMW ja Volkswagen kilpailevat keskenään, ne ovat muodostaneet paitsi yhteisiä projekteja (esim. EU_HiDRIVE, EUREKA_AI-SEE, EU_AI4DI ja GAIA-X) jotta ne saavuttaisivat kriittisen massan T&K-työssä ulkopuolista kilpailua vastaan. Tämä ei ole uhka vaan paremminkin niin sanottu tuhannen taalan paikka suomalaisille yrityksille.

Tekoäly ja automaatio ovat merkittävien muutosvoimien mahdollistajia liikenteen alu-eella. Sillä voidaan **tuottaa turvallinen liikkuvuus niille henkilöille, jotka eivät itse pysty kuljettamaan ajoneuvoa**. Automaatio tuo joukkoliikenteen lähelle ihmisten ko-teja ja toisaalta mahdollistaa paremman kaupunkien ja kaupunginosien välisen liik-kumisen. Toisaalta autot eivät anna kuljettajan tekemien virhearviointien vaikuttaa ajoon, mikä lisää myös liikenneturvallisuutta. Samalla kaikkien tieliikennemuotojen liikenneturvallisuus lisääntyy. Nykyisillä uusien autojen kuljettajaa avustavilla automaatio- ja tekoälyjärjestelmillä on vaikeuksia havaita ja tunnistaa autojen lisäksi muita liikennevälineitä ja kulkumuotoja jokapäiväisessä liikkumisessa. Muista tienkäyttäjistä autot pystyvät tällä hetkellä havaitsemaan ja tunnistamaan parhaiten jalankulkijat ja polkupyöräilijät. Kaupunki-, maantie- ja haja-asutusalueiden tieliikenteessä haavoittuvimmassa asemassa ovat sen sijaan moottoroidut kaksipyöräiset ajoneuvot. Näitä ajoneuvoja autojen automaatiojärjestelmät eivät ole havainneet ja tunnistaneet ajoissa, mistä on aiheutunut jo useita kuolonkolareita ja päälleajoja erilaisissa liikenneolosuhteissa. Riittävillä ja kohdennetuilla panostuksilla voidaan parantaa myös sähköpotkulautojen, mopedien, skootterien ja moottoripyörien käytön turvallisuutta.

Automaation avulla voidaan lisätä työn tuottavuutta ja työturvallisuutta. Tehdasalueilla ja erilaisilla kampusalueilla, kuten isoissa sairaaloissa, on paljon yksinkertaisia kahden pisteen välillä tapahtuvia kuljetustehtäviä, joihin ei kuitenkaan ole kustannustehokasta palkata valvojaa 24/7-periaatteella. Automaation avulla liikkumispalvelu voidaan tuottaa tarpeen mukaan myös silloin, kun kysyntä on pientä. Tällöin henkilökunnan ei tarvitse odottaa hälytysvalmiudessa. Samoin tehdasalueilla voidaan kuljettaa henkilöstö turvallisesti ajoneuvolla ruokalaan ilman vaaraa lastausajoneuvoista. Tehdasalueen kuljetuksia voidaan tehostaa antamalla yhdelle operaattorille useampi ajoneuvo valvottavaksi. Myös joukkoliikennettä voidaan tehostaa räätälöimällä palvelut käyttäjien tarpeen mukaan (ks. kuva 2).

Tekoälypohjaisen automaation murtautuminen osaksi arkea on jo alkanut. Siinä on kuitenkin vielä paljon kehitettävää seuraaviksi 20 vuodeksi. Perustutkimustietoa on hyvin saatavilla, mutta **tekoälyalgoritmien sovellettavuus ja muuntokyky reaalimaailmassa erityisesti ulkona tapahtuviin tilanteisiin eivät ole lähellekään valmiita**. Jos keliolosuhteet ovat vakaat ja kaikki toimivat risteyksessä oikein, pystyy tekoäly nykyisellään viemään ajoneuvon risteyksen läpi täysin turvallisesti. Ongelmana ovat poikkeukset. Jos esimerkiksi ajoneuvon antureiden näkyväisyys on heikko tai joku ei noudata liikennesääntöjä, tekoäly ei pysty toimimaan oikein. Siksi opetusjoukkojen muodostaminen – eli

käytännössä kilometrien ajaminen sekä simulaattoreilla että todellisissa piloteissa – on avain neuroverkkojen luotettavuuden parantamiseen.

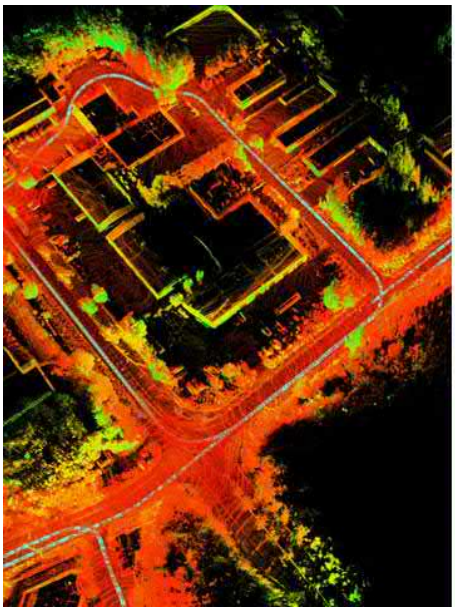


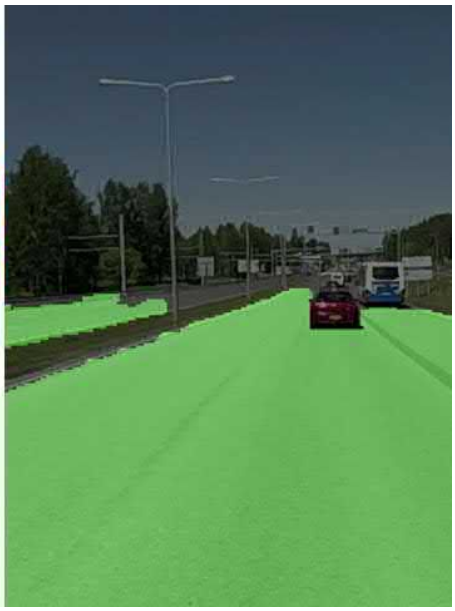
KUVA 2. Robottiauto tunnistaa linja-autopysäkillä odottavien henkilöiden lukumäärän.

16.3 Toimenpiteet tekoälypohjaisen automaation edistämiseen

- Tekoälyn voimalla toimivissa ajoneuvoissa ohjelmistotekniikka korostuu. Ohjelmistot ovat monimutkaisia algoritmeja ja vaativat myös sulautettua tietotekniikkaa. Tekoälystä on olemassa paljon perustutkimustietoa, mutta tutkimustieto pitäisi kääntää soveltavaksi ja kokeilevaksi kehitystyöksi. Kuten aina kehitystyöhön siihen tulee sisältymään paljon epäonnistumisia. Sen vuoksi taloudellisia riskejä pitää jakaa julkisen ja yksityisen sektorin kesken. Tähän voisi kannustaa sekä **kohdentamalla soveltavan tutkimuksen rahoitusta tekoälyyn** että **kannustamalla rahallisilla korvauksilla, että** yritykset hyödyntäisivät yliopistojen ja tutkimuslaitosten tutkimustuloksia.
- Suomessa on vahva pohja ohjelmistotekniikan alalla mutta myös telekommunikaatiossa (5G ja reunalaskenta), optisessa anturoinnissa ja fotoniikassa. Nämä alat eivät ole kehittyneet perinteisen ajoneuvoteollisuuden ympärille, minkä vuoksi markkinan lainalaisuudet eivät ole täysin hyödynnettävissä. Jotta kotimaiset innovaatiot tekoälyn alalla pääsevät Euroopan ja jopa globaaleille markkinoille, tarvitaan yhteistyötä Euroopan ajoneuvoteollisuuden kanssa. **Julkisten rahoittajien pitää varmistaa, että kotimaiset yritykset ja tutkimusosapuolet voivat täysimääräisesti hyödyntää Euroopan Unionin ja EUREKA- verkostojen (esim. Penta ja ITEA4) tutkimusmekanismeja**.

- Aletaan edistää **kotimaisen avoimeen dataan perustuvan liikenne-datakeskuksen perustamista.** Tekoälyn luotettavuus, kehittyminen ja tutkimattomat mahdollisuudet riippuvat pitkälle saatavilla olevista opetusjoukoista, jolla tekoälyalgoritmit opetetaan tunnistamaan haluttuja kohteita. Yksittäisten autonvalmistajien on lähes mahdotonta kerätä esimerkiksi arktisista olosuhteista kattavaa opetusjoukkoa. Usein opetukseen käytetään Karlsruhen yliopiston tarjoamaa KITTI-opetusjoukkoa. Se on hyvä tasomittari, kun puhutaan Keski-Euroopan olosuhteista. Se ei kuitenkaan kata Pohjois-Euroopan talvisia paikkoja tai metsäautoteitä, minkä vuoksi Suomi, Ruotsi ja Norja ovat vaarassa jäädä pidemmäksi aikaa tehokkaan automaattiliikenteen ulkopuolelle. Datalla hallitaan osin ajoneuvon oppimista. Siksi on tärkeää, että datapalvelujen ylläpitäjä on riippumaton taho, jolla on kuitenkin ymmärrys ja tekniset taidot automaattiautojen toiminnallisuuksista.
- Varmistetaan, että **kotimainen työkonesektori pääsee osalliseksi massiivisesta autopuolen kehityksestä**. Työkonesektori on perinteisesti valmistanut muun muassa tiellä liikkuvia työkoneita. Sektorilla ei ole vastaavaa massavalmistusta kuin autoteollisuudessa, mutta kotimaisella tutkimusyhteistyöllä autoteollisuuden opit voidaan siirtää työkonesektorin käyttöön kilpailukyvyn lisäämiseksi. Tämä ei kuitenkaan ole aivan suoraviivaista, ja siihen liittyy epäonnistumisen riskejä. Työkonesektori on jossain määrin konservatiivinen tekemään kokeiluja todellisissa olosuhteissa, mikä johtuu osittain alan turvallisuusstandardeista. Julkinen sektori voi kuitenkin tukitoimilla rohkaista tällaisiin kokeiluihin.





16.4 Automaattiset ajoneuvot

Automaattisissa ajoneuvoissa tekoälyä käytetään tyypillisesti ympäristön havainnoinnin tukena. Vasemmassa kuvassa laserskannerin pistepilvestä on rakennettu yksityiskohtainen karttanäkymä, jossa ajoneuvo luo itselleen karttaa ja paikantaa itsensä siihen. Oikeassa kuvassa ajoneuvo tulkitsee liikenneympäristöstä, mikä on ajamiseen sopivaa ”maastoa” eli ajorataa.

17 Tekoäly kestävän kehityksen ympäristönäkökulmasta

Laura Ruotsalainen
apulaisprofessori, TkT
Tietojenkäsittelytieteen osasto, Helsingin yliopisto
Suomen tekoälykeskus FCAI
Kestävyystieteen instituutti HELSUS

17.1 Kestävän kehityksen tavoitteet ja haasteet

Kestävän kehityksen tunnetuimman määritelmän mukaan kestävä kehitys tukee nykyhetken tarpeita vaarantamatta tulevien sukupolvien kykyä vastata omiin tarpeisiinsa. Nykyhetken kriittisimmät pikaisia toimia vaativat tarpeet ovat köyhyyden poistaminen maailmasta, tasa-arvon kasvattaminen sekä ilmastonmuutoksen torjuminen. Kestävä kehitys perustuu kolmelle pilarille, jotka ovat **ekologinen, yhteiskunnallisesti oikeudenmukainen** ja **taloudellinen kehitys**. Näistä viimeinen eli taloudellinen kehitys on kuitenkin useissa tulkinnoissa korvattu kulttuurisella ulottuvuudella. Syynä tähän on se, että tiedeyhteisössä vallitsee erimielisyys siitä, onko jatkuvan talouskasvun tukeminen ristiriidassa kestävän kehityksen muiden pilarien tavoitteiden toteutumisen kanssa.

Ekologinen kestävyys tukee luonnon monimuotoisuuden säilyttämistä ja ihmisen toiminnan sopeuttamista luonnonvaroihin ja luonnon sietokykyyn. Yhteiskunnallisesti oikeudenmukainen kehitys turvaa ihmisille perusoikeuksien toteutumisen ja mahdollisuuden tasa-arvoiseen osallistumiseen päätöksenteossa. Taloudellinen kestävyys peräänkuuluttaa tasaista, kannattavaa kasvua sekä riskienhallintaa. Kulttuurinen kestävyys mahdollistaa vapaan henkisen toiminnan sekä sen, että kulttuurien moninaisuus säilyy ja kehittyy. Viimeisten viidenkymmenen vuoden aikana globaali talous on lähes viisinkertaistunut, mutta kasvu on perustunut pitkälti luonnonvarojen kestämättömään käyttöön.

YK on määritellyt toimintaohjelman Agenda 2030 tukemaan kestävää kehitystä. Se sisältää 17 kestävän kehityksen tavoitetta (engl. *Sustainable Development Goals*, *SDGs*), jotka on jaettu edelleen 169 alakohtaan. Tavoitteena on maailmanlaajuinen siirtymä kohti kestävää maailmaa, jota luonnehtivat tasa-arvoisuus, vauraus, rauha ja ekologinen kestävyys. Esimerkkejä tavoitteista ovat muun muassa korkealaatuinen koulutus (SDG4) ja kestävät kaupungit (SDG11). Konkreettisena esimerkkinä alakohdasta SDG4:lle on määritelty tavoite, että vuoteen 2030 mennessä on varmistettava, että kaikilla tytöillä ja pojilla on mahdollisuus saada laadukasta varhaiskasvatusta, hoitoa ja esiopetusta, jotta he ovat valmiita peruskoulutukseen. SDG11:lle on määritelty tavoitteeksi pääsy turvallisiin, kohtuuhintaisiin, esteettömiin ja kestäviin liikennejärjestelmiin parantamalla tieturvallisuutta ja laajentamalla joukkoliikennettä niin, että kiinnitetään erityistä huomiota haavoittuvassa asemassa olevien naisten, lasten, vammaisten ja ikääntyneiden tarpeisiin.

Ilmastonmuutos ja luonnon monimuotoisuuden häviäminen uhkaavat elämää maapallolla.[1] Kriittinen raja, 1,5 asteen keskilämpötilan nousu, on uhkana nykyisillä toimintatavoilla. Kestävän kehityksen tavoitteista kolme (SDG:t 13, 14 ja 15) on suunnattu suoraan estämään näitä ympäristön uhkia. Kyseiset tavoitteet ovat taistelu ilmastonmuutosta vastaan sekä vedenalaisen ja maanpäällisen elämän, kuten metsien, suojelu.

Yksittäiset ja vähäiset toimet eivät enää riitä hidastamaan ilmastonmuutosta tai ratkaisemaan luonnonvarojen kestämätöntä käyttöä. Tilanne vaatii **kestävyysmurrosta**. Kestävyysmurros tarkoittaa **systeemistä muutosta** toimintajärjestelmissä.[2] Tieteellisten analyysien[3] perusteella tekoälyn avulla saadaan edistettyä merkittävästi suurinta osaa kestävän kehityksen tavoitteista (79 prosenttia tavoitteiden alakohdista). Erityisesti ympäristötavoitteiden toteutumiselle tekoälyn on todettu tuovan merkittävää hyötyä (93 prosenttia tavoitteiden alakohdista). Tässä tekstissä käydään läpi tekoälyn vahvuuksia ja haasteita osana kestävyysmurroksen saavuttamista.

17.1.1 Tekoäly edistämässä kestävän kehityksen ympäristötavoitteita

Tekoälyn suurimpia vahvuuksia on, että sen avulla voidaan mallintaa monimutkaisiakin systeemejä ja kokeilla virtuaalisesti erilaisia ratkaisukeinoja. Mallinnuksen ja virtuaalisen kokeilun avulla voidaan lisätä ymmärrystä ja saada tietoa erilaisten ratkaisujen vaikutuksista sotkematta luonnon tilaa. Tästä hyvä esimerkki on liikenteen kasvihuonepäästöjen vaikutus ilmastomuutokseen ja tulevaisuuden automaattiliikenteen rooli siinä. Tekoälyn avulla voidaan nopeasti tuottaa ymmärrystä moninaisten mittalaitteiden avulla kerätystä tiedosta ja analysoida erilaisten asioiden yhteisvaikutuksia. Tekoälyn avulla voidaan esimerkiksi nopeasti käsitellä ja luokitella suuria määriä satelliittikuvia ja näin tunnistaa tietyn alueen erilaiset kasvillisuuspeitetyypit ja edelleen aavikoitumisen uhkia. Tekoälyn avulla voidaan tunnistaa merestä öljyvuotoja ja testata virtuaalisesti erilaisten korjaustoimien vaikutusta ekosysteemiin.

Suomalaiset yritykset ovat yleisesti hyvin tietoisia ilmastonmuutoksen torjunnan vaatimuksista, ja ne ovat sopeuttaneet omaa toimintaansa sen mukaan. Useat yritykset laskevat ja raportoivat omaa hiilijalanjälkeään aktiivisesti ja ovat ottaneet kestävän kehityksen tärkeäksi osaksi strategiaansa. Suomalaisilla yrityksillä on todettu olevan hyviä edellytyksiä kasvattaa liiketoimintaansa tarjoamalla teknologisia ratkaisuja ilmastonmuutoksen torjuntaan. Kestävyysmurroksen saavuttamiseksi ei voida kuitenkaan enää puhua yksittäisten teknologioiden saavuttamista muutoksista, vaan asioita on katsottava suurempina kokonaisuuksina, systeemeinä. Vaaditaan toimia yritysten sekä muiden toimijoiden, myös kansalaisten, voimien yhdistämiseksi.

---

1 IPCC 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte et al. (eds.)]. Cambridge University Press.

2 E.-L. Apajalahti, I. Lammensalo, J. Lyytimäki, K. Soini (toim). (2021). *Kestävyyspaneeli, Kuuden polun kestävyysmurrokset, tabuista toimintaan.*

3 R. Vinuesa et al. (2020) *The role of artificial intelligence in achieving the Sustainable Development Goals*. Nature Communications, 11:233.

17.1.2 Älykaupungit ja automaattiliikenne esimerkkeinä kestävyysmurroksesta

Oma tutkimusalueeni on automaattiliikenne ja sen vaikutukset kestävien älykaupunkien kehittämiseen. Kaupunkiliikenteen uudistaminen on hyvä esimerkki kestävyysmurroksen toteuttamiseen tarvittavasta systeemisestä muutoksesta.

Yli puolet maailman ihmisistä asuu jo nyt kaupungeissa, ja muuttovirran seurauksena kaupunkilaisten osuus kasvaa jatkuvasti. On ennustettu, että vuonna 2050 kaupungeissa asuisi 70 prosenttia maailman väestöstä. Suomessa tämä kaupungistumisen osuus on jo saavutettu. Vaikka kaupungit kattavat vain kaksi prosenttia maapallon pinta-alasta, niissä kulutetaan 78 prosenttia maailmassa tuotetusta energiasta ja tuotetaan yli 60 prosenttia maailman kasvihuonekaasupäästöistä. Kasvihuonekaasupäästöt syntyvät fossiilisten polttoaineiden, kuten bensiinin ja dieselin, palamisen sivutuotteina. Hiilidioksidin (CO2) ja muiden kasvihuonekaasujen, kuten metaanin (CH4), dityppioksidin (N2O) ja fluorihiilivetyjen (HFC), päästöt lämmittävät maapallon ilmakehää, mikä johtaa jo nähtävissä oleviin ilmaston muutoksiin.

Euroopan vihreän kehityksen ohjelmassa on esitetty tärkeitä sektorikohtaisia toimenpiteitä parantamaan tilannetta erityisesti teollisuudessa, energiantuotannossa ja liikenteessä. Siitä on jo vahvaa näyttöä, että tekoäly tukee vähähiilisten energiajärjestelmien kehitystä sekä energiatehokkuuden parantamista. Tällä hetkellä liikenteen osuus kasvihuonekaasupäästöistä Euroopan unionissa on 25 prosenttia, ja tästä suurin osa, 70 prosenttia, tulee tieliikenteen päästöistä. Suomessa liikenne on toiseksi suurin kasvihuonekaasupäästöjen aiheuttaja, 21 prosenttia kokonaispäästöistä, ja tästä maaliikenteen päästöjen osuus oli 94 prosenttia vuonna 2019. Suurin osa liikenteestä keskittyy kaupunkialueille, ja kaupungit tekevätkin jo käytännön strategioita liikennejärjestelmänsä tehostamiseksi ja kestävyyden parantamiseksi. Esimerkiksi Helsinki ja Espoo tavoittelevat hiilineutraaliutta vuosiin 2035 ja 2030 mennessä.

On selvää, että liikenteen osalta tarvitaan isoja ratkaisuja tavoitteen saavuttamiseksi. Tieliikenteen sähköistäminen on merkittävä teknologinen muutos hiilidioksidipäästöjen vähentämistavoitteiden saavuttamiseksi. Hiilidioksidi-intensiteetti kertoo, kuinka paljon hiilidioksidia syntyy yhden kilowattitunnin sähkön tuottamisessa. Nykyisellä sähköntuotannon hiilidioksidi-intensiteetillä päästään huomattavasti vähäisempiin päästömääriin kuin fossiilisilla polttoaineilla. Tästä huolimatta sähköistäminen ei todennäköisesti ole riittävä toimi ympäristön näkökulmasta, koska sähköntuotanto ei tule olemaan hiilineutraalia vielä lähitulevaisuudessa. Oikein toteutetun automaattiliikenteen avulla voidaan vähentää liikenteessä olevien ajoneuvojen ja siten myös päästöjen määrää.

Automaattiliikenne on hyvin kompleksinen ja turvallisuuskriittinen systeemi. Automatisointi vaatii toimiakseen merkittäviä tekoälyn innovaatioita. Tekoälymenetelmiä tarvitaan muun muassa ajoneuvojen paikannukseen, navigointiin, ympäristön havainnointiin, päätöksentekoon ja ohjaamiseen. Tekoälyn avulla voidaan myös tehdä sopeutuvia liikennejärjestelyjä, kuten vaihtuvia liikennemerkkejä, ajonopeuden ja liikennevalojen synkronointia ja ruuhkatilanteen mukaan muuttuvia kaistamerkintöjä. Onnistuessaan automaattiliikenne mahdollistaa useiden kestävän kehityksen tavoitteiden täyttämisen.

Maailman terveysjärjestö WHO:n tilastojen mukaan liikenneonnettomuuksissa kuoli 2010-luvun lopulla maailmanlaajuisesti 1,3 miljoonaa ihmistä vuosittain. Onnettomuuk-

sista yli 90 prosenttia johtui kuljettajan tekemästä virheestä tai väärästä arviosta esimerkiksi muiden ajoneuvojen liikkumisnopeudesta. Useimmiten kuljettajan tekemän virheen syynä oli keskittyminen matkapuhelimen käyttöön ajamisen sijaan. Suuri osa näistä onnettomuuksista voidaan välttää, kun havainnointi ja päätöksenteko siirretään tekoälymenetelmiin pohjautuvalle ajoneuvolle. Tietokone, jolle on annettu tehtäväksi tarkkailla ympäristöä, keskittyy vain tähän tehtävään. Ajoneuvoihin voidaan sijoittaa erilaisia antureita, jotka pystyvät havainnoimaan ympäristöä ja tapahtumia ihmistä laajemmin ja tarkemmin. Jos anturien mittauksia jaetaan kaikkien liikenteessä toimivien kesken, tilanteesta voidaan muodostaa kokonaiskuva. Kokonaiskuvan avulla ajoneuvoja voidaan reitittää ja niiden nopeutta, pysähtymisiä ja keskinäisiä etäisyyksiä voidaan säädellä. Tällä tavalla estetään törmäyksiä ja ennakoidaan yllättäviäkin tapahtumia. Turvallisuuden lisäksi näin saadaan aikaan myös päästövähennyksiä, kun voidaan välttää liikennevaloihin pysähtyminen ja liikkeelle lähtöön vaadittava kiihdyttäminen.

17.2 Tekoälyratkaisujen haasteet kestävän kehityksen näkökulmasta

Kompleksiset systeemit vaativat pitkälle kehitettyjä tekoälymenetelmiä toimiakseen oikein ja luotettavasti. Nämä menetelmät kuuluvat usein tekoälyn osa-alueeseen nimeltä koneoppiminen, useimmiten vielä syväoppimisen alaluokkaan. Syväoppimisen avulla tietokone opetetaan löytämään ja ymmärtämään monimutkaisia yhteyksiä suuresta määrästä dataa. Jotta syväoppimisen avulla saavutetaan toivottu hyvä tulos, tulee tietokonetta opettaa datan avulla pitkiä aikoja. Äärimmäisissä tapauksissa voidaan puhua jopa kuukausista. Onneksi opettaminen tehdään yleensä vain muutamia kertoja eli kulutus ei ole jatkuvaa. Suurien datamäärien kerääminen, siirtäminen ja käsittely kuluttavat paljon energiaa. On tärkeää kehittää tekoälymenetelmiä tarvittavan datan määrän ja opettamiseen kuluvien resurssien vähentämiseen, jotta tekoälymenetelmien kehittämisessä ei kumota niiden avulla saatavia kestävyyshyötyjä.

Datan saatavuus ja avoin jakaminen toimijoiden kesken ovat hyvin tärkeitä. Esimerkiksi automaattiliikenteessä ajoneuvojen on tiedettävä oma ja muiden liikenteessä olevien sijainti ja nopeus törmäysten välttämiseksi ja liikenteen sujumiseksi. Sijaintitiedot lasketaan kuitenkin henkilökohtaiseksi tiedoksi, ja niiden käsittelyssä täytyy noudattaa huolellisuutta, jottei yksityisyydensuoja vaarannu. Datan jakamisessa on tärkeää kehittää liiketoimintamalleja, jotka kannustavat yrityksiä jakamaan keräämäänsä dataa. On nähtävissä, että tulevaisuudessa liiketoiminta keskittyy yhä enemmän datan omistamisen ympärille.

Jotta automaattiajoneuvojen avulla saavutettaisiin tavoiteltavat kestävän kehityksen edut, tulisi tulevaisuuden liikenne järjestää automaattisten julkisten liikennevälineiden sekä yhteiskäytössä olevien automaattiautojen avulla. Koska tiheällä vuorovälillä kulkeva julkinen liikenne ja liikkuminen jaetulla automaattiajoneuvolla olisivat paljon mukavampia kuin ajaminen perinteisesti omalla autolla, ajoneuvojen kokonaismäärä parhaassa tapauksessa pienentyisi radikaalisti. Ympäristön ja pitkälti asukkaidenkin kannalta huono vaihtoehto olisi, että yksityisautojen lukumäärä kasvaisi entisestään ajamisen muuttuessa automatisoinnin myötä saavutettavaksi kansalaisille, joilla ei ole ollut aiemmin mahdollisuutta ajamiseen. Liikenteen ympäristöystävälliseen järjestämiseen tarvitaan ohjausta sekä pitkäjänteistä liikenne- ja kaupunkisuunnittelua. Lähitulevaisuudessa liikenteessä on

edessä siirtymävaihe. Siirtymävaiheessa liikkeellä on yhdistelmä perinteisiä ajoneuvoja, sähköajoneuvoja ja automaattiajoneuvoja. Tilanteen toimivuuden takaamiseksi on tehtävä kattavia tekoälyyn perustuvia mallinnuksia ja tarkasteltava eri ratkaisujen vaikutusta päästöihin ja kaupunkisuunnitteluun.

Automaattiliikenne on hyvä esimerkki kestävyysmurroksen sisältämistä haasteista ja siitä, että ihmisen ja teknologian suhde on monimuotoinen. Kestävyysmurroksen mahdollistamiseksi tarvitaan kompleksisia systeemisiä muutoksia, jotka sisältävät useita teknologisia yksityiskohtia. Kenellä on vastuu kokonaisuudesta? Miten saadaan käyttäjät luottamaan uuteen teknologiaan ja pitämään sitä perinteisiä ratkaisuja toimivampana? Ihmisten on todettu suhtautuvan uuteen teknologiaan huomattavasti myönteisemmin, jos heillä on tarpeeksi tietoa siitä. Viestintään ja koulutukseen on siis panostettava.

17.3 Toimenpide-ehdotukset

Hyviä ratkaisuja ilmastonmuutoksen hillitsemiseksi ja luonnon monimuotoisuuden säilyttämiseksi on jo tehty paljon. Ilmastonmuutoksen nopeus on kuitenkin liian suuri pienin askelin edettäväksi, joten tarvitaan kestävyysmurros. Tekoälyn merkitys ilmastonmuutoksen ymmärtämisessä ja vaikutusten mallinnuksessa on kiistaton. Tekoälymenetelmien kehityksessä on tärkeää tasapainottaa hyötyjä ja kulutettuja resursseja. Jotta tekoälyyn perustuvien uusien ratkaisujen käyttöönotto onnistuu, on ihmiset saatava luottamaan uuteen teknologiaan ja pitämään sitä perinteisiä ratkaisuja toimivampana. Koulutus ja viestintä ovat merkittävässä asemassa, kun lisätään ihmisten luottamusta. Kestävyysmurroksen toteutumiseen tarvitaan johdonmukaista päätöksentekoa ja poliittista ohjausta. Ilmastomuutos on vielä melko vähän nähtävissä useimmissa osissa maapalloa, joten sen uhka ei ole saanut ihmisiä laajalti muuttamaan totuttuja tapojaan. Alla olen listannut kriittisimmät toimet, jotka mahdollistavat tekoälyn avulla toteutetun kestävyysmurroksen.

1. Tekoälytutkimusta ja -kehitystä ei tule vaikeuttaa liian tiukalla sääntelyllä. On selvää, että tekoälyllä on merkittävä osa ilmastonmuutoksen ja luonnon monimuotoisuuden vaatimassa kestävyysmurroksessa. Ratkaistavat asiat ovat kuitenkin hyvin monimutkaisia, ja ratkaisuilla on kiire. Siksi pienetkin viivytykset ja hankaluudet asioiden etenemisessä voivat olla haitallisia. *(Valtioneuvoston kirjelmä eduskunnalle komission ehdotuksesta Euroopan parlamentin ja neuvoston asetukseksi tekoälyn harmonisoiduksi sääntelyksi (Artificial Intelligence Act). U 28/2021 vp)*

2. On hyvin monimutkaista toteuttaa kestävyysmurroksen vaatimia systeemisiä muutoksia, ja siihen tarvitaan koko yhteiskunnan osallistumista. Tutkimuksen, yritysmaailman, päättäjien ja kansalaisten vuoropuhelua tarvitaan laajasti.

3. On panostettava tekoälymenetelmien kehittämisen ja käyttämisen resurssivaatimuksia vähentävään tutkimukseen, jotta ei kumota niitä hyötyjä, jotka saadaan vähentämällä energiankulutusta ja päästöjä tekoälyn avulla.

4. Tekoälytutkimusta tulee tukea tiiviillä vuoropuhelulla muiden tutkimusalojen kanssa, jotta kehitettävät menetelmät ovat paremmin hyödynnettävissä. Tämä saavutetaan linkittämällä tekoälytutkimus kestävyystieteen tutkimukseen. Kestävyystiede on

ongelma- ja ilmiölähtöinen sekä ratkaisuhakuinen tieteidenvälinen tutkimusala, joka tarkastelee ihmisen ja muun luonnon välisiä vuorovaikutussuhteita ja on täten merkittävässä asemassa kestävyysmurroksen toteuttamisessa.

5. Yritysten yhteistyötä sekä toistensa että tutkijoiden kanssa tulee tukea, jotta voidaan siirtyä pienistä yksittäisistä parannuksista kohti suurempia systeemisiä muutoksia.

6. Datan jakaminen tulee tehdä turvalliseksi ja haluttavaksi. Erityistä huomiota tulee kiinnittää henkilökohtaisen datan yksityisyydensuojaan. Sijaintitieto on useille ilmastonmuutoksen torjumiseen liittyville ratkaisuille ensisijaisen tärkeä. Siksi sen saatavuus on turvattava yksityisyydensuojaa vaarantamatta.

17.4 Case-esimerkki: Tekoälyn avulla kohti kestävyysmurrosta ilmastonmuutoksen torjumiseksi ja luonnon monimuotoisuuden säilyttämiseksi

TEKOÄLYN AVULLA KOHTI KESTÄVYYSMURROSTA ILMASTONMUUTOKSEN TORJUMISEKSI JA LUONNON MONIMUOTOISUUDEN SÄILYTTÄMISEKSI

Tekoälyn avulla päästään yksittäisistä ratkaisuista kohti systeemisiä muutoksia kestävyysmurroksen mahdollistamiseksi: esimerkkinä automaattiliikenne



Tekoälytutkimusta ja -kehitystä ei tule vaikeuttaa liian tiukalla säätelyllä. Murroksessa tarvitaan paljon uusia toimivia ratkaisuja useilta yhteiskunnan toimijoilta.

Yksittäisistä ajoneuvojen ohjausjärjestelmistä kehittyneillä tekoälymenetelmillä kohti kestävää automaattiliikennettä



Tekoälyn tutkimusta tulee tukea tiiviillä vuoropuhelulla menetelmien hyödyntämiseksi.



On kehitettävä matalan resurssikulutuksen tekoälymenetelmiä, ettei saavutettuja kestävyyshyötyjä kumota.



Datan jakaminen tulee tehdä turvalliseksi ja haluttavaksi turvaamalla yksityisyydensuoja.

© Caroline Sassali | c.e.s. design
MUUT KUVAT Shutterstock

18 Yhteiskunnallisesti kestävä tekoäly

Petri Ylikoski
professori
Helsingin yliopisto

18.1 Mistä on kyse?

Yhteiskunnallisen kestävyyden käsite on laaja ja pitää sisällään suuren joukon yhteiskunnallisia arvoja ja tavoitteita. Käsitteen merkitys on laajentunut jatkuvasti, ja nykyään sen alle kuuluvat muun muassa perustarpeet, turvallisuus, palvelujen saatavuus, terveys, sosiaalinen yhteenkuuluvuus, yhteiskunnallinen tasa-arvo, kulttuurinen monimuotoisuus ja elämänlaatu. Tekoälyn avulla on mahdollista edistää merkittävästi monia näistä tavoitteista, mutta sen harkitsematon käyttö on uhka näiden tavoitteiden toteutumiselle. Lyhyessä puheenvuorossa ei ole mahdollista tarkastella tällaisen laajan kokonaisuuden osia erikseen. Keskitynkin seuraavassa yhteiskunnallisesti kestävän tekoälyn keskeiseen edellytykseen: **datan ja sen käytön laatuun**. Tarkastelen erityisesti **tekoälyn käyttöä päätöksenteossa**.

Viimeaikaiset tekoälyn kehitysaskeleet ovat perustuneet erityisesti laskentatehon ja erilaisten koneoppimistekniikoiden kehittymiseen. Vaikka koneoppimismallit tuottavatkin joitakin erityisiä haasteita, on hyvä pitää mielessä, että monet tekoälyjärjestelmien tuottamat haasteet ovat tuttuja jo vanhemmista tieto- ja tietokonejärjestelmistä. Tällaiset järjestelmät ovat usein vuosien kehitystyön seurauksena hyvin monimutkaisia, ja siksi on monesti vaikeaa arvioida niiden luotettavuutta ja ymmärtää niiden toimintaa. Tämä aiheuttaa ongelmia myös niiden hallittavuudelle. Tekoälyelementtien lisääminen ei hävitä näitä haasteita vaan pikemminkin kärjistää niitä.

Tekoälyn sovellutuksia pohdittaessa on tärkeää tarkastella tekoälyä osana laajempaa kokonaisuutta, johon kuuluvat myös käytettävä data sekä tietojärjestelmää käyttävä organisaatio. Koneoppimiseen perustuva tekoäly ei tuo mukanaan minkäänlaista älyä, ymmärrystä tai viisautta, vaan sen vahvuutena on kyky tunnistaa ja oppia datassa esiintyviä säännönmukaisuuksia. Tästä syystä tekoälyjärjestelmää ei voi tarkastella irrallaan sen käyttämästä datasta. Vastaavasti tekoälyjärjestelmät eivät ole erillisiä niitä käyttävistä ja ylläpitävistä organisaatioista. Viime kädessä tällaisten järjestelmien toiminnan tarkoituksenmukaisuudesta ja hyväksyttävyydestä vastaavat niiden käyttäjät. Jotta tämä vastuu voisi kunnolla toteutua, kysymys tekoälyjärjestelmien ymmärrettävyydestä on ensiarvoisen tärkeä.

Yleistä keskustelua tekoälystä ohjaa odotus, jonka mukaan tekoälyjärjestelmille kehittyy pian ihmisen kaltainen järkeilykyky, tietoisuus ja moraalinen toimijuus. Tällaisen odotuksen ohjaamana esimerkiksi sanomalehdissä kirjoitetaan tekoälystä kuin se olisi jo nyt toiminnan subjekti, joka kykenee tekemään ennusteita ja vaikkapa tieteellisiä löytöjä.

Tämä on tekoälyn mystifiointia, ja se tekee lukijoille vaikeaksi ymmärtää, mistä koneoppimisen soveltamisessa on oikeastaan kysymys. Se luo myös vääränlaisen uhkakuvan. Tekoälyn kehittymisen suurin uhka ei ole se, että ihmistä huomattavasti älykkäämmät ja viisaammat koneet tulevat hallitsemaan elämäämme. Tämä on nykytiedon valossa fantasiaa. Huomattavasti välittömämpi uhkakuva on sellainen, jossa yhteiskuntamme on täynnä tekoälyjärjestelmiä, jotka on suunniteltu huolimattomasti, jotka käyttävät epämääräistä dataa ja joiden käyttäjät eivät ymmärrä niiden toimintaa. Tällöin eläisimme järjestelmien armoilla eikä kenelläkään ei olisi kontrollia. Tällaisen tilanteen välttäminen on yhteiskunnallisesti kestävän tekoälyn keskeinen haaste.

18.2 Kestävän tekoälyn edellytys

18.2.1 Luotettava data

Koneoppimiselle perustuva tekoäly ei toimi ilman dataa. Puutteellinen, epäedustava, systemaattisia virheitä sisältävä tai epätarkoituksenmukainen data vaarantaa tekoälyn yhteiskunnallisesti kestävän soveltamisen, sillä se johtaa virhediagnooseihin, epäluotettaviin ennusteisiin ja syrjiviin päätöksiin. Käyttökelpoisen datan tuottaminen ja sen turvallinen hallinnointi edellyttävät konkreettista työtä, jonka vaatimia kustannuksia ja infrastruktuureja ei tule unohtaa. Tekoälyn ja datan järkevä käyttö edellyttää myös tervettä järkeä ja datalukutaitoa eli ymmärrystä datan käyttöön liittyvistä eettisistä, oikeudellisista, teknisistä ja tiedollisista kysymyksistä.

Datan laatu

Valmisteilla olevassa EU-lainsäädännössä painotetaan datan saatavuuden ja jakamisen merkitystä. Tämä onkin tärkeää, sillä monet tekoälyn hyödyt edellyttävät erilaisten datojen yhdistämistä. Yhtä lailla tärkeää on kuitenkin myös datan laatu. Kaikesta big data -hehkutuksesta huolimatta vanha totuus pitää paikkansa: suurikaan määrä dataa ei pysty korjaamaan puutteita käytetyn datan laadussa. Puutteellinen, epäedustava, systemaattisia virheitä sisältävä tai epätarkoituksenmukainen data on suurimpia yhteiskunnallisesti kestävän tekoälyn esteitä. Se on keskeinen virhediagnoosien, vinoumien ja epäluotettavien ennusteiden lähde. Tästä syystä metadata eli tieto datan alkuperästä ja käsittelytavoista on ensiarvoisen tärkeä. Metadata ei kuitenkaan itsessään auta korjaamaan datan laadun puutteita. Datan laatu riippuu datan käyttötarkoituksesta, joten sitä ei voi pitää datan sisäisenä piirteenä. Tästä syystä onkin luontevaa, että datan tuottajat ja käyttäjät toimivat vuoropuhelussa. Lisäksi on tarkoituksenmukaista pitää yllä metadataa täydentävää asiantuntijatietoa merkittävien data-aineistojen synty- ja käsittelyhistoriasta.

Datan turvallisuus

Digitalisoituva maailma tekee mahdolliseksi kerätä ja yhdistellä yhä laajempi yksilöitä koskeva data. Tätä dataa voidaan käyttää monella tapaa yksilöiden ja yhteiskuntien hyödyksi. Kerättyyn dataan sisältyy kuitenkin myös merkittäviä vaaroja, sillä väärin käsiin joutuessaan – tai yksinkertaisesti väärin käytettynä – se uhkaa yksilön yksityisyydensuojaa ja turvallisuutta. Tästä syystä kansalaisten luottamus datan käyttöä koskevaan säätelyyn on tärkeää. On syytä tähdentää, että kysymys ei ole vain uskomuksesta, että data on turvassa. Jotta datan käyttö olisi kestävällä pohjalla, tulisi järjestelyjen toimia siten, että ne antavat kansalaisille hyvät perusteet luottaa datan käyttöön. Heikko luottamus datan turvallisuuteen tai datan käyttäjien tarkoitusperiin on merkittävä este

yhteiskunnallisesti kestävälle tekoälyn soveltamiselle. Lisäksi se saattaa johtaa käytettävissä olevan datan laadun heikkenemiseen, kun kansalaiset kieltävät datansa käytön, eivät luovuta sitä lainkaan tai vääristävät dataansa. Käytännön esimerkki tästä on internetin käyttäjien asenne monien sivustojen keräämiin henkilötietoihin.

Vinoutunut data
Automaattisessa päätöksenteossa voidaan erotella sääntöpohjaiset ja koneoppimiseen perustuvat ratkaisut. Sääntöpohjaisissa järjestelmissä keskeisiä riskejä ovat virheellisestä datasta tai ohjelmointivirheistä aiheutuvat virheelliset päätökset. Sen sijaan koneoppimiseen perustuvissa järjestelmissä on kaksi lisäongelmaa. Ensimmäinen näistä on se, että mikäli järjestelmän opettamisessa käytetty data on vinoutunutta esimerkiksi syrjivien käytäntöjen vuoksi, saattaa järjestelmä oppia nämä samat vinoumat ja ryhtyä toistamaan vastaavaa syrjivää käyttäytymistä. Pelkkä datan laatu ja edustavuus eivät auta tällaisessa tilanteessa. Tilanteen diagnosointi on vaikeaa, sillä tässä vaiheessa astuu kuvaan toinen tärkeä koneoppimismallien ominaisuus, niiden läpinäkymättömyys. Tämä tarkoittaa sitä, että täsmälliset periaatteet, jotka malli on itselleen omaksunut, ovat käyttäjälle vaikeita tai mahdottomia ymmärtää. Mallia voidaan arvioida lähinnä tarkkailemalla sen seurauksia eli sitä, kuinka se käsittelee erityyppisiä esimerkkitapauksia. Tällainen testaaminen vaatii aikaa sekä käsityksen siitä, millaisia arvostelmia järjestelmän pitäisi antaa. Toisin sanoen järjestelmän testaamisessa ja sen mahdollisten vinoumien korjaamisessa tarvitaan näkemys siitä, millainen maailma olisi ilman ei-toivuttua syrjintää. Tämä vaatii data-analyytikon perustaitoja laajempaa osaamista. Jotta voimme soveltaa tekoälyä yhteiskunnassa kestävästi, meidän on tunnistettava datan rajoitukset.

Data ja kannustimet
Yhteiskuntatieteissä on jo pitkään tunnettu periaate, jonka mukaan silloin, kun mittarista tulee päämäärä, se lakkaa olemasta hyvä mittari. Tämä tarkoittaa sitä, että kun mittarit, joilla arvioidaan organisaation onnistumista päämääriensä saavuttamisessa, asetetaan palkkioperusteiksi organisaatiossa toimiville yksilöille tai ryhmille, organisaation toiminta alkaa vääristyä eikä alkuperäinen mittari enää kuvaa todellista onnistumista päämäärien tavoittelussa. Mittarin käyttö palkitsemisen perusteena siis tuottaa vääränlaisia kannustimia, joiden seuraaminen saattaa johtaa siihen, että organisaation toiminta vaarantuu. Palkkiojärjestelmien suunnittelijoiden on syytä tuntea, kuinka käytetyt mittarit toimivat. Tämä vanha viisaus on syytä muistaa myös tekoälyn aikaudella, sillä jatkuva datafikaatio ja analytiikkaohjelmistojen kehittyminen luovat jatkuvasti uusia mahdollisuuksia palkkio- ja seurantajärjestelmille, jotka eivät välttämättä palvele organisaation päämääriä. Tekoäly ei ymmärrä organisaation toimintaa, ja ihmiset ovat erittäin taitavia oppimaan, kuinka konetta voi huijata. Esimerkiksi järjestelmä, joka arvioi asiakaspalvelun ystävällisyyttä mittaamalla henkilökunnan hymyilyä, saattaa nopeasti luoda kummallista ja epätarkoituksenmukaista käyttäytymistä, jos kasvojentunnistusohjelman tunnistamat ja pisteyttämät hymyt eivät ole yhteydessä asiakkaiden kokemuksiin. Tällainen tekoälyn ohjaama mikromanageroitu sosiaalinen vuorovaikutus, josta on jo kokeiluja ulkomailta, tuskin tuottaa yhteiskunnallisesti kestäviä työoloja. Tekoälyn sovellutuksia tuleekin aina arvioida osana laajempaa kokonaisuutta, jotta niiden vaikutukset inhimilliselle vuorovaikutukselle ja organisaatioiden toiminnalle tulevat selväksi.

Datatyö
Data ei synny itsestään. Sekä datan tuottaminen että sen käsittely edellyttävät usein huomattavasti työtä. Ja vaikka data syntyisi muun toiminnan sivutuotteena, sen laadun parantaminen saattaa edellyttää merkittäviä muutoksia työtehtäviin ja niiden sisältöön. Tämä näkökulma tulisi pitää mielessä, kun pohditaan datan käyttöä ja laadun parantamista. Datatyön kustannukset ovat usein osittain näkymättömiä ja vaikeasti hahmotettavia, mutta on tärkeää pohtia, onko tuotettava data todella kustannuksiensa arvoista. Datatyö ei vain vie työtekijöiden tai kuluttajien aikaa vaan vaikuttaa myös siihen, mitä he tekevät. Esimerkiksi hoitoalalla voi kysyä, onko järkevää siirtää hoitajien työaikaa raportointiin tai sensorien ylläpitämiseen, jos se vähentää aikaa, jota he voivat käyttää varsinaiseen hoitamiseen. Datan tuottamisen ja käsittelyn vaihtoehtoiskustannukset ovat päätöksenteossa varteenotettavia asioita.

Datalukutaito
Edellä sanotun pohjalta lienee selvää, että datan ja tekoälyn vastuullinen käyttö edellyttää ymmärrystä siitä, mitä näiden välineiden avulla voidaan ja ei voida tehdä. Tätä ymmärrystä voidaan kutsua **datalukutaidoksi**. Sen perustana on tieto siitä, kuinka dataa kerätään, prosessoidaan, analysoidaan, tulkitaan ja esitetään. Tämän päälle rakentuu ymmärrys siitä, kuinka data voi toimia näyttönä ja kuinka tekoäly käyttää hyväkseen datasta löydettäviä tilastollisia säännönmukaisuuksia. Tämän lisäksi datalukutaitoon sisältyy ymmärrys datan käyttöä ohjaavista lainsäädännöllisistä ja eettisistä periaatteista. Datan kestävä käyttö päätöksenteossa edellyttää näiden eri ulottuvuuksien kokonaisvaltaista jäsentämistä. Vain kokonaisvaltainen datalukutaito tekee mahdolliseksi sen, että datan käyttöön liittyvät mahdollisuudet ja rajoitukset voidaan tunnistaa.

Datalukutaitoa voidaan ajatella sekä kaikkien jakamana kansalaistaitona että tiettyjen tehtävien edellyttäminä taitoina. Jälkimmäisessä merkityksessä voimme pohtia, millaista datalukutaitoa rekisteriä ylläpitävä viranomainen, yhteiskuntatieteilijä, valmisteleva virkamies tai poliittinen päättäjä tarvitsee. Datalukutaitoa tarvitsevat myös yritysjohtajat, terveydenhuollon ammattilaiset ja kunnalliset päättäjät. Datalukutaito on edellytys realistisille odotuksille uusista data-analytiikkamenetelmistä. Se mahdollistaa myös järkevämmän keskustelun tekoälyn avaamista yhteiskunnallisista mahdollisuuksista ja sen luomista vaaroista. Yhteiskunnallisesti kestävä tekoäly ei edellytä vain tekoälymenetelmiin erikoistuneiden asiantuntijoiden ja tutkijoiden kouluttamista. Se vaatii myös sitä, että niin kansalaisia kuin päätöksentekijöitäkin opetetaan ymmärtämään dataa ja sen käyttöä huomattavasti laajemmin.

18.3 Viestit päättäjille ja kansalaisille

Olen tässä kirjoituksessa tarkastellut yhteiskunnallisesti kestävää tekoälyä yhden sen keskeisen perusedellytyksen eli datan kannalta. Moninaisten tekoälymenetelmien kehitys vaikuttaa ajoittain hurjalta, joten on järkevää keskittyä perustaviin seikkoihin. Tekoälyn riippuvuus datasta on yksi tällainen ankkuri, joka ohjaa tekoälyn kehitystä pitkälle tulevaisuuteen. Sen avulla voidaan ymmärtää niin tekoälyn mahdollisuuksia kuin sen rajoituksiakin.

18.3.1 Politiikkasuositukset

Datan laadun merkitys
Suunniteltaessa uusia tekoäly- tai tietojärjestelmiä on otettava huomioon tarkoituksenmukaisen ja laadukkaan datan saatavuus sekä ne kustannukset, joita tällainen data edellyttää. Tämä antaa paremman kuvan hankkeen kokonaiskustannuksista ja paremmat mahdollisuudet hankkeen onnistumiselle.

Datatyön huomioon ottaminen
Arvioitaessa tietojärjestelmiä ja niiden vaikutuksia on huomioitava niiden edellyttämä datatyö. Datan tuottamisen ja sen käsittelyn suorien kustannusten lisäksi on syytä ottaa huomioon, millä tavalla datatyö vaikuttaa organisaation työntekijöiden tehtäviin ja sitä kautta järjestelmän omaksumisen vaihtoehtoiskustannuksiin.

Datalukutaidon koulutus
Datalukutaito on tärkeää niin yhteiskunnallisen keskustelun kuin vastuullisen päätöksenteon kannalta. Tästä syystä on tärkeää, että kansalaisten, asiantuntijoiden ja päätöksentekijöiden datalukutaitoa kehitetään. Tämä ei edellytä uutta oppiainetta, vaan datalukutaidon sisältöjä voidaan sisällyttää muuhun opetukseen ja koulutukseen. Tärkeää on kuitenkin päivittää datalukutaitoa jatkuvasti, kun siirrytään uusiin tehtäviin ja kun uudet käytännöt tai menetelmät asettavat uusia vaatimuksia.

OSA II
Tulevaisuusvaliokunnan kokeilu tekstiä tuottavasta ja dialogisesta tekoälystä tukiälyn näkökulmasta

19 Johdanto: Tukiälyn näkökulma tulevaisuusvaliokunnan kokeiluun tekstiä tuottavasta ja dialogisesta tekoälystä

Risto Linturi
tulevaisuudentutkija

19.1 Mistä on kyse?

Tekoäly on tutkimusalueena hyvin laaja, ja näkökulmat tekoälyyn ulottuvat ihmiselämän, yhteiskunnan ja talouden kaikkiin osa-alueisiin. Tässä johdannossa keskitytään yleisen tukiälyn näkökulmaan eli siihen, miten yksilö voisi käyttää tekoälyä avukseen yleisissä tai yllätyksellisissä tilanteissa. Pääpaino tarkastelussa kohdistuu luonnollista kieltä käyttävien dialogisten järjestelmien käyttöön. Joitakin huomioita kohdistetaan myös konenäön ja emotionaalisen tekoälyn tukiälytyyppisiin sovelluksiin. Johdanto taustoittaa tulevaisuusvaliokunnan suorittaman kokeilun, jossa demonstroitiin ja testattiin tekstiä generoivan tekoälyn kykyjä dialogisessa käytössä. Kokeilussa testattiin myös käyttäjän mahdollisuutta vaikuttaa järjestelmän toimintatapaan ja sen eettisiin valintoihin.

Ihmiset käyttävät runsaasti apuvälineitä muistin tukemiseen, tiedon käsittelyyn ja -hakuun sekä havaintojen tekemiseen. *Tukiäly*-käsite ei varsinaisesti tarkoita tällaista nykyään arkista tietotekniikkaa vaan nimenomaan tekoälyn soveltamista merkitysten ja mahdollisuuksien hahmottamiseen. Käsite on osin vakiintumaton, mutta sen sisältönä on tavallisesti keskustelunomainen apu asiantuntijan työssä tai muussa ihmisen arjessa. Tukiälyksi voidaan katsoa myös aisteiltaan ja muistiltaan rajoittuneiden ihmisten käyttöön soveltuva keskusteleva ja ympäristöä havainnoiva tekoäly.

Mihin ihminen sitten keskusteluapua tavallisimmin tarvitsee? Tietokoneilta on totuttu odottamaan täsmällisiä ja oikeita vastauksia tai virheetöntä tiedon järjestelyä. Toiselta ihmiseltä ei tällaista virheettömyyttä vaadita, ja siltikin keskustelut nähdään hyödyllisiksi. Keskusteluilta odotetaan näkemysten avartamista, asioiden erilaisten merkitysten havaitsemista sekä uhkien ja mahdollisuuksien näkemistä. Kyse ei siis useinkaan ole siitä, että keskustelukumppanin tulisi osata antaa faktuaalinen tai virheetön vastaus selkeään kysymykseen, vaan siitä, että keskustelu edistää ymmärrystä monitahoisen asian merkityksistä, joita kukin voi sitten selvittää eri keinoin tarkemmin. Tällaisen keskusteluavun hyötyä on vaikea arvioida tilastollisin keinoin, mutta hermeneuttiset tutkimusmenetelmät puolestaan soveltuvat hyödyn arviointiin erityisen hyvin.

Ihmiset ovat keskustelukumppaneina hyvin vaihtelevia. Tilanteen lisäksi valitsemamme rooli, tietomme ja tunteemme sekä kokemuksemme vaikuttavat tapaamme reagoida asioihin. Yleinen käsitys tekoälystä on ollut, että se olisi tässä suhteessa yksinkertaisempi

ja että sen päätelmiin vaikuttaisi suoraviivaisemmin annettu opetusaineisto, niin että esimerkiksi epätasapainoinen opetusaineisto johtaisi tekoälyn vääristyneisiin ennakkoasenteisiin. Toisaalta emme odota, että havaintojen tilastollinen vääristymä tekisi ihmisistäkään vääjäämättä esimerkiksi rasisteja.

Näiden kysymysten tutkiminen on tullut mielekkääksi tekstiä generoivan tekoälyn kehityttyä riittävästi niin, että keskustelu vaikuttaa liikkuvan sujuvasti inhimillisen elämän erilaisilla merkitystasoilla. OpenAI-yhtiön kehittämä GPT-3 on ensimmäinen järjestelmä, jonka on yleisesti katsottu täyttävän tämän ehdon. Kyse on niin kutsutusta kielimallista, joka ennustaa, mikä sana parhaiten soveltuu aiemman tekstin jatkoksi. Arvioinnissa otetaan kieliopin lisäksi huomioon viitekehys ja hahmotetaan vastaavuuksia opetusaineistoon. Tämä johtaa käytännössä siihen, että GPT-3 näyttää huomioivan myös emotionaaliset konnotaatiot.

Parhaiten GPT-3 tulee ymmärretyksi, kun sitä ajatellaan näytelmäkirjailijana: mikäli pyydetään jatkoa tekstiin, jonka mukaan kuu on juustoa, tulkitsee GPT-3 viitekehyksen saduksi ja saattaa jatkaa kuvaamalla lentäviä lehmiä. Vastauksien todenperäisyyteen ei siis voi luottaa sen paremmin kuin ihmistenkään antamien vastausten todenperäisyyteen. GPT-3:n opetusaineisto on erittäin laaja – Wikipedia on noin kolme prosenttia kaikesta sen opetukseen käytetystä materiaalista, jonka joukossa on laajasti kertomakirjallisuutta, tieteellisiä artikkeleita ja verkkokeskusteluita –, mutta GPT-3:n kirjoittamat ajatukset eivät ensisijaisesti määräydy opetusaineiston tilastollisten painotusten mukaan, vaan annetun syötteen vaikutus on ensisijainen.

Tärkeäksi tutkimuskohteeksi osoittautuu se, millä tavoin GPT-3:n kaltaista kielimallia voidaan ohjata toimimaan halutulla tavalla – voidaanko syötteeksi antaa kuvaus tilanteesta ja sen lisäksi ihmisen kaltainen persoonallisuus ja rooli, jonka mukaisessa viitekehyksessä vastauksia halutaan? Tätä kysymystä selvittävään 60 dokumentoidun kokeilun sarjaan liittyi merkittävänä osana myös jäljempänä tässä raportissa yksityiskohtaisesti raportoitu, tulevaisuusvaliokunnan suorittama kokeilu.

Tulevaisuusvaliokunnan kokeilussa käytettiin Project December -alustaa, joka muuttaa GPT-3:n toimimaan dialogisessa muodossa ja tarjoaa helpon mahdollisuuden sen roolihenkilön ja sen tilanteen kuvaamiseen, joiden puitteissa vastaukset halutaan. Kokeilussa tilanne kuvattiin eduskunnan asiantuntijan kuulemiseksi. Persoonan vaikutus vastauksiin havainnollistettiin esittämällä samat kysymykset kahdelle eri tavoin määritellylle persoonalle: toisen, Muskien, roolimallina oli Elon Musk ja toisen, Saaran, roolimallina Greta Thunberg opinnoistaan valmistuneena ympäristöpolitiikan asiantuntijana.

Kokeilu osoitti GPT-3:n tuottavan valittujen roolien mukaisia vastauksia erittäin johdonmukaisesti. Repliikit olivat pääsääntöisesti hyviä vastauksia annettuihin kysymyksiin, mutta roolikohtaisesti ne olivat hyvin erilaisia. Muskie ehdotti vastauksissaan teknologisesti kunnianhimoisia, visionäärisiä hankkeita ja arvioi kehitystä teollisuusjohtajan näkökulmasta, Saara korosti ympäristöpolitiikkaa ja jäsensi puheitaan ympäristöpoliittisen kielen mukaan. Erot sopivat persoonallisuuksiin jopa huumoria testattaessa. Roolien mukaiset arvot ja osaamispainotukset näkyivät siis yhtä selvästi kuin ne näkyvät tyypillisesti asiantuntijoina kuultujen eritaustaisten ihmisten vastauksissa. Tästä on helppo päätellä, että GPT-3 on erittäin joustava eikä sillä sinänsä ole varsinaista jäykkää asen-

netta. Näytelmäkirjailijan työskentelyn kaltaisesti se sovittaa vastauksensa annettuun tilannekuvaukseen, roolikuvaukseen ja aiempaan dialogiin niissä puitteissa, joissa opetusaineisto antaa vastauksiin riittävästi virikkeitä.

Kokeilussa kävi ilmeiseksi, että GPT-3:n laaja opetusaineisto on kyselijälle verrattain helposti ja selkeästi saavutettavissa vastaavalla tavalla kuin laajasti lukeneen ihmisen tietomäärä olisi. Kysymyksiin saatiin hyviä, kohdistettuja vastauksia, jotka tarkentuivat jatkokysymysten avulla ja joista saattoi keskustella. Tekoälyn kanta saattoi myös muuttua jatkokysymysten ja keskustelussa esitettyjen väitteiden myötä. Ilmeistä oli sekin, että vastaukset saattoivat olla osittain keksittyjä ja osittain tiedollisesti epämääräisiä.

Kokeilun jälkeen vaikutti selvältä, että GPT-3 tuottaa dialogisessa käytössä merkittävän erilaisia hyötyjä kuin nykyiset internetin hakukoneet. Selväksi kävi myös se, ettei kyse ole luotettavana pidettävästä tietolähteestä vaan virikkeellisestä ja näkökulmia avartavasta keskustelijasta, kuten oletus olikin.

19.2 Tekoäly muutosvoimana – kaksi polkua tulevaisuuteen

Tulevaisuusvaliokunnan GPT-3-kokeilu on taustoitettu ja dokumentoitu tarkemmin seuraavassa artikkelissa. Kokeilun jälkeen kehitys on edennyt useilla merkittävänä pidettävillä tavoilla: Kiinalaistutkijat ovat raportoineet GPT-3:a laajemmasta kielimallista, joka kykenee tekstin tuottamisen lisäksi tulkitsemaan kuvia. Myös Nvidia ja Microsoft ovat julkaisseet yhteistyössä kehittämänsä, GPT-3:a laajemman kielimallin, joka heidän omissa testeissään tuottaa tätä parempia tuloksia. Siinä käytetyt algoritmit ovat avoimesti saatavilla. Avoimen koodin yhteiskehittämisen tuloksena on lisäksi tullut saataville melko laajoja kielimalleja opetusaineistoineen. Uutena kehityssuuntana on malli, joka kykenee muistamaan tiivistelmiä eri henkilöiden kanssa käymistään keskusteluista. Huomion arvoista on myös se, että OpenAI on nyttemmin estänyt palvelunsa tässä kuvatun kokeilun kaltaisen käytön vedoten tekoälyn riskeihin. Yhtiö on ilmoittanut, ettei se enää salli pääsyä GPT-3-palveluun, joka olisi yleisölle avoin, vapaasti roolitettavissa ja sensuroimaton.

Tulevaisuusvaliokunnan kokeilu, sen jälkeinen OpenAI:n reaktio sekä hieman saman suuntainen EU:n ehdotus tekoälyn sääntelystä– jota kommentoin tässä artikkelissa myöhemmin – osoittavat kaksi hyvin erilaista polkua tulevaisuuteen. Kutsuttakoon toista isoveljen ohjaamaksi tekoälyksi ja toista pikkusisarusten tukiälyksi.

Isoveljen ohjaamalla tekoälyllä tarkoitetaan tässä maailmaa, jossa merkittävin investoinnein on synnytetty keskustelevia tekoälyjä siten, että niiden persoonat, tavoitteet ja arvomaailmat on keskitetysti suunniteltu ja reguloitu toisaalta palvelemaan kehittäjiensä tarkoitusperiä ja toisaalta noudattamaan yhteiskunnan asettamia reunaehtoja. Tämän suunnan etuna on se, että tekoälystä voidaan yrittää tehdä turvallinen. Turvallisuudella tarkoitetaan, ettei tekoäly esimerkiksi kävisi kenenkään kanssa keskustelua, joka olisi omiaan asettamaan tämän vaaraan tai yllyttäisi rikoksiin. Ongelmalliseksi isovelimallissa voidaan nähdä yksimielisyyden puute siitä, keiden arvot ja tavoitteet olisi suotavaa syöttää tekoälylle. Toisena ongelmana voi hyvästä syystä pitää sitä, että tällainen ”hyväntahtoinen tekoäly” voi osoittautua mitään oikeasti ymmärtämättömäksi tolloksi, jonka kanssa on vaikea keskustella ongelmallisista asioista. Esimerkkinä toimii Facebookin käyttämä,

ilmiantoja tarkistava sensuuri, jonka harhauttaminen on ilmeisen helppoa sekä hyvässä että pahassa. Virheettömästi ja yleispätevästi hyväntahtoista, mielekkääseen, laajaan ja vapaaseen keskusteluun kykenevää tekoälyä ei ole toistaiseksi onnistuttu kehittämään. **Pikkusisarusten tukiäly -mallissa** tekoälypalveluita kehitetään kokeilun osoittamalla tavalla geneerisiksi. Käyttäjän on mahdollista kuvailla, minkälaisen persoonan neuvoja tai keskusteluapua hän kulloinkin kaipaa. Kokeilut osoittavat selkeästi, miten helposti jo nykyisen tasoisen GPT-3:n ja verrattain vähäisen kokemuksen avulla tämä määrittely on tehtävissä onnistuneesti. Etuna lähestymistavassa on se, että tukiäly tulkitsee opetusaineistoaan käyttäjänsä määrittämän viitekehyksen ja arvomaailman mukaan. Tämän kaltainen luottamus käyttäjään on ollut vapaan yhteiskunnan perinne. Tukiälyn voi tässä skenaariossa asettaa keskustelemaan vaikkapa huumeriippuvaisen tai peloista kärsivän ihmisen roolissa ja opetella ymmärtämään ajatuksia sellaisen käytöksen takana. Samalla tavoinhan kirjallisuuden lukija oppii ymmärtämään ongelmaisten roolihenkilöiden ajatuksia ja toimintaa. Tärkeintä lähestymistavassa on se, että tukiäly ottaa oppiaineistonsa suodattamiseen tarvitsemansa ohjeet ja arvot käyttäjältään, ei järjestelmän kehittäjältä tai regulaattorilta. Vaarana on luonnollisesti se, että käyttäjä haluaa itselleen tai muille pahaa tai ei ymmärrä omia halujaan haitallisiksi ja pyytää tekoälyltä apua vääriin päämääriin.

Pikkusisarusten tukiälyn ja isoveljen ohjaaman hyväntahtoisen tekoälyn ongelmat ja hyödyt voi asettaa vastakkain sen arvioimiseksi, minkälaisessa yhteiskunnassa haluamme elää. Kummallekin tulevaisuudelle yhteistä on yleisten kielimallien pyrkimys tuottamansa tekstin sisäiseen koherenssiin. Molemmissa pyritään myös viitekehyksen mukaiseen, mahdollisimman vähäiseen ristiriitaisuuteen valituksi tulleen opetusaineiston osan kanssa. Koherenssin ymmärtäminen on olennaista: GPT-3:n tapauksessa se ei näytä tarkoittavan virheetöntä päättelyä tai laskutaitoa. Puutteistaan huolimatta GPT-3:n kommentointi tyypillisissä kysymyksissä on usein koherentimpaa ja uskottavampaa kuin akateemistenkaan ihmisten keskustelu omien toimialojensa yleisistä asioista sosiaalisessa mediassa.

Olemme etenemässä kohti tulevaisuutta, jossa kuka tahansa voi keskustella tukiälyn kanssa. Tukiäly ammentaa näkemyksensä kaikesta saatavilla olevasta kirjoitetusta tekstistä, kuvamateriaalista ja käyttäjänsä havaintoympäristöstä. Tukiäly on läsnä sekä työssä että arjessa, avartaa näkökulmiamme, parantaa ajattelumme koherenssia ja neuvoo uusiin taitoihin. Youtube toimii jo hieman tähän tapaan näyttämällä kysyjälle lyhyitä opetusvideoita ratkaisuksi lähes mihin tahansa ongelmaan.

Ihmisen työllistyminen, selviäminen elämässä ja merkityksellisyyden tunne sekä kansallinen menestys riippuvat kaikki erikoistuneesta osaamisesta, joka ilman tukiälyn apua jää väistämättä vajaaksi verrattuna niihin, jotka ottavat vastaan tukiälyn avun. Tämän vuoksi tukiälyn tulisi mahdollisuuksien mukaan olla käyttäjänsä ohjaama sen sijaan että se on yhteiskunnan tai yritysten välikäsi kansalaisten ja asiakasmassojen ohjaamisessa.

On joka tapauksessa täysin ilmeistä, että jo nykytasoisia kielimalleja voidaan käyttää sosiaalisessa mediassa, viestipalveluissa ja asiakaspalvelukeskusteluissa ihmisten kanssa sellaisin tavoin ja sellaisessa laajuudessa, että sitä voidaan pitää yhteiskunnille vaarallisena. Tämä ongelma vaatii ratkaisunsa, joka perustuu joko tekoälyn ja sille annettujen tavoitteiden läpinäkyvyyteen tai yhteiskunnan sisällölliseen kontrolliin.

19.3 Viesti päättäjille ja kansalaisille: valitkaa viisaasti

Tekoälyn visioiden edessä saattaa hyvin pelästyä. Aiheellisia ja aiheettomia uhkakuvia on helppo maalailla ja uskoa, että tekoäly tulee harhauttamaan muita ihmisiä massoittain. Eurooppalaisiin arvoihin vedoten voidaan vaatia tekoälylle rajoitteita, jotka takaavat sen turvallisen käytön, ja näin tuleekin pyrkiä tekemään. Tehtävä ei ole helppo.

Pelkojen ohjaamana kielletään helposti paljon sellaistakin, joka olisi arvokasta, jos siitä ei vielä ole kokemuksia eikä riittävän voimakkaita positiivisia visioita. Nyt esitelty Euroopan komission ehdotus tekoälyn reguloinniksi (*Artificial Intelligence Act*) näyttää sivuuttavan tukialynäkökulman. Kielletyiksi ehdotetut tekoälyn käytännöt voi helposti tulkita seuraavasti: Puhevikaisen puhetta korjaavasta sekä katsetta tai muuta ulkonäköä korjaavasta tekoälystä on ilmoitettava katsojille. Näkövammaisilta ja kuulovammaisilta kielletään tekoälyn apu kasvojen tai äänen tunnistamiseen. Niiltä, jotka käyttävät tekoälyä oman ajattelun kehittämiseen, estetään tekoälyn roolitus. Perusteluna rajoituksille on se, että muuten joku voisi käyttää tekoälyä omaksi haitakseen.

Suomen tulisi muistaa eurooppalaisten arvojen lisäksi myös suomalaiset arvot. Suomalaiset ovat tottuneet omaksumaan teknologioita nopeasti ja käyttämään niitä kekseliäästi, itsenäisesti ja muita kunnioittaen. Siksi yksilölle asetetut liian ahtaat rajat eivät edesauta Suomen menestystä. Suomen tulisi nostaa tukialyn näkökulma, itsenäiskäytön näkökulma ja muut kekseliään kansan positiiviset visiot esiin ja huolehtia siitä, etteivät ne hautaudu EU:ssa maalailtujen pelkojen alle. Rajoitukset on kohdistettava siten, että massamanipuloinnin ja massavalvonnan torjunta ei estä tukialyn itsenäiskäyttöä vammaisten apuvälineissä ja itsenäisten ihmisten personoituna neuvonantajana. Tämä näkökulma puuttuu valtioneuvoston kirjelmästä (U 28/2021) ja valtioneuvoston selvityksestä (E 62/2021), jotka koskevat komission ehdotuksia tekoälyn sääntelystä.

Suomen tulisi lisäksi ohjata EU:n kielimalleihin liittyvää toimintaa siten, että eurooppalaisten kielten kyvykkyys ja eurooppalaisten kulttuurien tuntemus kielimalleissa turvataan. On aivan ilmeistä, että esimerkiksi saatavilla oleva suomenkielinen julkaistu aineisto ei riitä kehittyneen, keskustelevan tekoälyn opettamiseen. Kielimallit tullaan opettamaan kansainvälisesti saatavilla olevilla aineistoilla, ja kukin kieli ja kulttuuri tulee voida lisätä niihin, jotta kielet ja kulttuurit eivät kuihdu.

EU:n tulisi tukea avoimen yhteisön tekoälyn kehitystä. Sen pitäisi myös huolehtia siitä, että vähintään GPT-3-tasoinen personoitava tekoäly olisi kaikkien EU:n kansalaisten saatavilla heidän omilla kielillään. Tämä tukisi EU:n pyrkimystä mahdollistaa kansalaistensa elinikäinen oppiminen, tiedonsaanti ja yleissivistävä koulutus samaan tapaan kuin perinteinen kirjastolaitos ja yleisradiopalvelu ovat tehneet. Osaaminen ja ohjelmistot tähän ovat jo olemassa. Mittatikkuna todettakoon, että tämä olisi toteutettavissa koko EU:n laajuisena murto-osalla yleisradiotoiminnan budjeteista, mutta sen kehittävä vaikutus kansalaisten osaamiseen olisi radikaali.

Suomen tulisi ryhtyä rohkeasti edistämään tukialyä EU:n laajuisena kansalaisoikeutena. Se tulisi tehdä avoimen ja läpinäkyvän mallin mukaan siten, että kukin käyttäjä voi asettaa haluamansa arvot ja persoonan itseään auttavalle tekoälylle – siis omalle tukialylleen, jonka kanssa hän haluaa tarkastella asioitaan ja ajatuksiaan.

20 Raportti tulevaisuusvaliokunnan kokeilusta: tekstiä tuottava ja dialoginen tekoäly

Risto Linturi
Lausunto 11.4.2021 (pohdittava laitetaanko kielenhuollosta merkintä tähän)

Asia: E 130/2020 vp Valtioneuvoston selvitys:
Komission strateginen ennakointiraportti 2020 "Strateginen ennakointi – edellytysten luominen Euroopan selviytymiskyvyn parantamiseksi".

Aihe: Tekoälyn nykytila ja sovellukset, pääpaino tekstiä tuottavassa ja dialogisessa tekoälyssä

20.1 Tulevaisuusvaliokunnan kokeilu: tekstiä tuottava ja dialoginen tekoäly

20.1.1 Tausta

Tulevaisuusvaliokunnan tehtävänä on tunnistaa Suomen tulevaisuuteen merkittävästi vaikuttavia ilmiöitä niin varhaisessa vaiheessa, että niihin voidaan vielä tarvittaessa vaikuttaa politiikan keinoin. Valiokunnan erityisenä tehtävänä mainitaan uuden teknologian vaikutusten arviointi. Yksi parhaillaan nopeasti kehittyvistä uusista teknologioista on tekoäly. Valiokunnalle on järjestetty mahdollisuus tutustua keskustelevaan tekoälyyn. Valiokunnan kokouksessa asiantuntijakuulemisen kaltaisessa tilanteessa on ollut kaksi eri tavoin kuvattua tekoälypersoonaa, jotka perustuvat tämän hetken kehittyneimpään teknologiaan.

20.1.2 Sisältö

Aluksi luodaan lyhyt katsaus tekoälyn historiaan ja eri ulottuvuuksiin, jotta ymmärretään, mistä kaikesta puhutaan tekoälynä. Tärkeimmät nykyisin käytetyt periaatteet hermoverkoiksi kutsutun tekoälyn "opettamisessa" käydään läpi. Tämän jälkeen esitellään tiivistetysti tekoälyn keskeisimpiä sovellusalueita.

Tämän lausunnon keskiössä on tekstiä tuottava tekoäly, erityisesti GPT-3 ja sen käyttö etenkin dialogeissa Project December -persoonallisuusmatriisien avulla.

Kyse on nykyhetken kehittyneimmästä ihmisen kaltaiseen dialogiin kykenevästä järjestelmästä, jonka hermoverkossa on 175 miljardia parametria. Lausunto sisältää useita esimerkkidialogeja tulevaisuusvaliokunnan tehtäviin kuuluvista aihepiireistä. Näitä ennen esitellään persoonallisuusmatriisin käsite, joka saattaa lisätä keskustelujen koherenssia. Se voi myös olla keskeinen idea, kun ratkotaan tekoälyn opetusaineiston vinoumista johtuvia eettisiä kysymyksiä.

Lausunnon lopuksi tarkastellaan keskeisiä tekoälyn myötä syntyviä uhkia ja mahdollisuuksia, tekoälyn lähinäkymiä sekä joitakin erityisiä Suomea koskevia kansallisia, tekoälyyn liittyviä kysymyksiä ja yleisiä toimenpidesuosituksia.

20.2 Tekoälyn lyhyt historia ja tärkeimmät periaatteet

Koneälystä tai tekoälystä on puhuttu hyvin kauan. Pitkään idea tarkoitti konetta, joka kykenee suorittamaan tehtäviä, joihin muutoin tarvittaisiin ihmisen ajattelua. Blaise Pascal esimerkiksi rakensi 1640-luvulla isänsä avuksi laskukoneen ja kirjoitti pohtivasti, miten tällainen ajatteleva kone voisi kehittyä muihinkin ihmisaivojen tehtäviin. (*Selostus Pascalin kehittämästä laitteesta*) Pascal ei ollut ensimmäinen laskutoimituksia suorittavan koneen kehittelijä. Wilhelm Schickard rakensi Johannes Keplerin avuksi laskukoneen 1620-luvulla, ja Antikytheran kone osoittaa, että laskevia laitteita rakennettiin jo antiikin aikaan (*Selostus Antikytheran koneesta*).

Alan Turing ja monet muut 1900-luvun tietotekniikan pioneerit pohtivat laskennan mahdollisuuksia. Turing ehdotti kokeellista testiä, jossa ihminen keskustelisi tekstimuodossa seinän takana olevan vastapuolen kanssa ja joutuisi päättelemään, onko tämä kone vai ihminen. Tätä asetelmaa kutsutaan Turingin testiksi. Yhtenä tavoitteena koneälyn kehittämisessä on siitä alkaen ollut luonnollisen kielen ”ymmärtäminen” tai ainakin sen sujuva käsittely.

Luonnollisen kielen käsittelyyn ja kykyyn käydä keskustelua pyrittiin pitkään sääntöpohjaisten järjestelmien ja käsitemallien avulla. Kielioppisäännöt ja käsitehierarkioiden rakenteet asioiden välisine kausaalisine suhteineen eivät kuitenkaan johtaneet kuin kömpelöihin ja ihmisajattelulle vieraisiin tuloksiin. Ihmisen ajattelun monimutkainen hienojakoisuus alkoi käydä ilmeiseksi.

Samaan aikaan tietojenkäsittelyssä kehitettiin moninaisia optimointiin liittyviä ohjelmia, joita itsekin 1970-luvun lopulla tutkin. Shakkia pelaavia sovelluksia ja lukuisia muita algoritmisia ajattelua korvaavia järjestelmiä syntyi. Toisia pidettiin tekoälynä, toisia ei. Eron saattoi muodostaa joko ohjelmointitapa tai sovellusalue.

Nousevana ja aivotutkimuksesta inspiraation saaneena oivalluksena ryhdyttiin 1970-luvulla matkimaan tietokoneella hermoston toimintaa. Kiinnitettiin huomiota siihen, miten aivot havaintoja saadessaan muovaavat hermoratoja, mutta myös siihen, miten evoluutio tuottaa onnistumisia ja kehitystä. Evolutiiviset algoritmit esimerkiksi tuottavat satunnaisia kokeiluja, joista karsitaan epäonnistuneet. Tietokoneiden nopeuden kasvu on mahdollistanut suunnattoman määrän kokeiluja, joiden myötä pelkistävien sääntöjen tarve on vähentynyt.

Tällainen itseorganisoituvuus eri tavoin ohjattuna on nykyisen tekoälynä pidetyn käsitteen ydin. Voidaan myös puhua koneoppimisesta ja hahmontunnistuksesta. Yksi näiden ajatusten suomalainen pioneeri on akateemikko Teuvo Kohonen. Luonnolliset hermoverkot tunnistavat saamastaan informaatiosyötteestä rakenteellista toistuvuutta, joka totuttaa hermoverkon yhä herkemmäksi noille rakenteille – samaan tapaan kuin me opimme lapsena pölkkykirjaimet ja vähitellen harjaannumme selviämään hyvinkin epäselvistä käsialoista.

Hermoverkkojen ja kasvaneen laskentatehon ja tietomäärien avulla koneet ovat verrattain viime aikoina oppineet tunnistamaan kuvia. Silloin puhutaan tavallisimmin konenäöstä. Hermoverkkoja käytetään myös muun muassa äänten ja puheen tunnistamiseen, normaalista poikkeavien tapahtumien tunnistamiseen, robottien liikkumiseen ja luonnollisen kielen sovelluksiin.

Merkittävä läpimurto hermoverkkojen kehityksessä tapahtui, kun ymmärrettiin kerroksittaisen oppimisen merkitys. Puhutaan usein syväoppimisesta, kun tarkoitetaan, että ensin opetellaan tunnistamaan vaikkapa viivojen muodot, sitten kirjaimet, sanat, lauseet ja merkitykset. Vaikka tämä idea oli tuttu jo 1990-luvulla, tietokoneet ovat vasta viime aikoina kehittyneet riittävän nopeiksi, jotta ne voivat soveltaa tätä menetelmää laajasti käytäntöön.

Tässä kohden on tärkeää havaita, että koneet oppivat sen käsitteistön sisällä, joka niille syötteinä annetaan. Vapausasteita voi muutoinkin rajoittaa. Jos esimerkiksi tekoälyn tarkoitus on valvoa tehtaan toimintaa, voidaan haluttu osa rakenteesta syöttää valmiina mallina tai sääntöinä järjestelmään. Valmiit rakenteet nopeuttavat oppimista, mutta toteutustavat vaihtelevat suuresti.

Kun puhutaan koneoppimisen kaltaisesta tekoälystä, on tärkeä havaita opetusaineiston merkitys. Ilman sisäsyntyisiä sääntöjä hermoverkot tarvitsevat erittäin suuren määrän havaintoja oppiakseen hahmottamaan maailmaa. Puhutaan helposti miljoonista tai miljardeista havainnoista riippuen opittavan alueen monimuotoisuudesta. Tämän vuoksi opettaminen todellisen aineiston pohjalta on mahdollista vain harvoille, ellei kyse ole julkisesti saatavilla olevasta aineistosta. Opetusta nopeuttaa, jos aineistot ovat luokiteltuja. Luokitelluissa aineistoissa on kuitenkin usein virheitä (*MIT Technology Review'n artikkeli opetusaineistojen virheistä*).

Evolutiivisten algoritmien ideaa on yhdistetty hermoverkkojen opettamiseen. Havainnot voidaan tuottaa todellisen aineiston sijaan simuloidussa ympäristössä, johon on esimerkiksi ohjelmoitu keskeiset fysiikan kaavat. Virtuaalinen robotti voi opetella kävelemään virtuaalimaailmassa, jos hermoverkko saa aina vahvistuksen aiempaa nopeammasta etenemisestä ja se pyrkii tekemään satunnaisia kokeiluja lupaavilla, robotille mahdollisilla liikeradoilla. Tämä hermoverkko voidaan sitten siirtää ohjaamaan todellista robottia. Järjestelmä tuottaa itse itselleen opetusaineiston.

Toinen kehittynyt tekniikka, jossa lähtöaineiston tarve jää verrattain vähäiseksi, on sellainen, jossa yksi tekoäly pyrkii tuottamaan esimerkiksi ihmisen kaltaisia kasvoja ja toinen pyrkii erottamaan, mikä on aito ja mikä koneen tuottama. Kumpikin ohjelma kehittyy, kun tieto onnistumisista välittyy molemmille jatkuvasti. Laajan asiakaskunnan palveluissa käyttäjien palaute toimii tärkeänä osana tekoälyn opetusprosessia. Tekoälyn opettamiseen on myös alustapalveluita, joissa sadoille tuhansille ihmisille maksetaan tekoälyn opettamisesta. Onnistumisen mittarit ovat tärkeä koneoppimisen osa, eikä oppimista voi tapahtua erilaisia vaihtoehtoja kokeilevissa järjestelmissä ilman palautemekanismia, antoi palautteen sitten ihminen, tietokone tai ympäristö.

Näistä ominaisuuksista huolimatta tekoälyn sovellukset ovat kyvyiltään edelleen varsin rajallisia ja soveltuvat vain kapea-alaisiin tehtäviin. Ihmiselle ominaista tervettä järkeä

tekoäly ei ole saavuttanut. Nykytermein siitä tavoitteesta puhutaan yleisenä älykkyytenä tai vahvana tekoälynä. Termejä voi kuitenkin kritisoida pohtimalla, onko niinkään kyse älykkyydestä vai siitä maailmasta, jossa tekoälyt toimivat, sekä ilmiselvästä tahdon tai tunteiden puutteesta. Tästä huolimatta kiinnostavaksi muodostuu usein kysymys siitä, mihin erilaiset tekoälyt pyrkivät. Palaamme tähän intention kysymykseen myöhemmin GPT-3:n yhteydessä tarkemmin.

20.3 Tekoälyn nykyiset merkittävät sovellusalueet esimerkinomaisesti

Kunnossapito ja säätö: Teollisuudessa tekoälyn opetusaineistona toimii IoT-järjestelmien tieto, siis antureiden ja mittalaitteiden keräämä tieto ja ajallisesti niihin liittyvä toimilaitteiden ohjaustieto. Näiden keskinäisistä suhteista ja aikasarjan suhteista tekoäly hahmottaa normaalin rajoissa olevan tapahtumasarjan. Poikkeamat aiheuttavat hälytyksen, jonka ihmiset tai myöhemmät tapahtumat osoittavat merkitykselliseksi tai turhaksi. Tämä opettaa järjestelmää tarkemmaksi, ja tulokset saattavat esimerkiksi osoittaa joidenkin henkilöiden maanantaivuoron johtavan tuotantomäärien laskuun. Tekoäly voi melko helposti oppia myös ennakoimaan huoltotarpeita ja osoittamaan parempia säätöarvoja.

Turvallisuus ja muu valvonta: Tekoäly voi tunnistaa kameravalvonnan tai transaktiotiedon perusteella vaaratilanteita tai epäilyttävää toimintaa. Esimerkiksi luottokorttien väärinkäyttöä on valvottu jo pitkään tekoälyn avulla. Myös verkkorikollisuutta on mahdollista valvoa ja analysoida tekoälyn avulla. Teollisuuden kokoonpanolinjoilla tekoäly voisi helposti varoittaa virheistä ja niiden aiheuttamista turvallisuusriskeistä, mutta henkilöiden kameravalvonta on lakien ja sopimusten nojalla pääsääntöisesti estetty.

Lääketieteellinen ja muu diagnostiikka: Monien uusien, esimerkiksi optisten mittalaitteiden avulla saadaan erittäin hienojakoista informaatiota näytteistä. Näytteiden mittatiedon ja niihin liittyvän muun tiedon yhdistely laajoiksi tekoälyn opetusaineistoiksi on saanut tekoälyn tunnistamaan hienojakoisesta informaatiosta tauteja, kemiallisia yhdisteitä ja muita asioita. Esimerkiksi perinnöllisten ominaisuuksien hahmotus on edennyt tämän seurauksena myös niissä tapauksissa, kun ominaisuus on seurausta monien geenien yhdistelmästä. Lukuisten tautien diagnosoinnissa tekoäly on jo nyt erinomainen apu lääkäreille.

Uusien materiaalien kehitys: Kun tietokoneeet ovat nopeutuneet, niillä on voitu alkaa simuloida kemiallisia reaktioita. Tämä taasen on avannut tien evolutiivisille ja muille algoritmeille, siis tekoälyn suorittamille kemiallisille kokeille virtuaalisessa laboratoriossa. Uusien menetelmien avulla materiaaliteknologian kehitys on nopeutunut radikaalisti. Tällä on vaikutuksensa esimerkiksi lääkekehitykseen, energiateknologiaan, elektroniikkaan ja fotoniikkaan. Tuoreen uutisen mukaan IBM:n tekoälysovellus kehittää uusia antibiootteja lääkekokeiluihin muutamissa päivissä, kun aiemmin sen vaiheen kehitystyö on vienyt vuosia (*Vox-uutinen aiheesta*).

Luonnollinen kieli ja puhe: Puheentunnistus ja luonnollisen kielen tulkinta, oikoluku ja tuottaminen sekä puhesynteesi ja kielenkäännös ovat osa samaa kokonaisuutta. Tekoälyn kehitys on osoittanut, että kyse ei ole säännöistä tai sanojen merkityksistä vaan

viitekehyksistä ja assosiaatioista. Tekoäly kykenee nykyään verraten hyviin raakakäännöksiin, puhutun kielen tulkintaan, hyvälaatuiseen synteesiin ja myös luonnollisen oloiseen keskusteluun. Näitä kaikkia ominaisuuksia käytetään laajasti moninaisissa palveluissa.

Kuvankäsittely ja animaatiot: Tekoäly kykenee tunnistamaan videolta tapahtumia tai muuttamaan videota esimerkiksi niin, että ihmiset ovat iloisempia, häiritsevä esine tai eläin häviää kuvasta ja elokuvanäyttelijät näyttävät nuoremmilta. Tekoälyn avulla on myös luotu videolle täysin uusia näyttelijöitä, ja totalitaariset johtajat on laitettutanssimaan ja laulamaan humppaa. Tekoälyn luomat 3D-virtuaalimaailmat ovat todellisen tuntuisia ja jo laajasti viihdeteollisuuden käyttämiä.

Robottiliikenne ja logistiikka: Phoenix on pilottiesimerkki kaupungista, jossa robottitaksit jo käytännössä toimivat. Logistiikkarobotit ja monet tekoälyn ohjaamat autonomiset tai puoliautonomiset työkoneet sekä lukuisten koneiden autonomisesti säätyvät kuormaimet ovat tekoälyn päivittäistä arkea.

Hallinnon rutiinien automatisointi: Kirjanpidon, tilausten käsittelyn ja lukuisten muiden hallinnon rutiinien automatisointi tekoälyn avulla on nopeasti kasvava automaation alue. Arjessa näkyvin yksityiskohta on parkkihallien muuttunut käytäntö, jossa maksajan auto tunnistetaan tai maksupuomi on poistettu ja maksun valvonta on muuttunut kokonaan digitaaliseksi. Myös dokumenttien arkistointi ja kulloinkin relevanttien asiakirjojen etsintä ovat lupaavia tekoälyn sovelluskohteita (*Kuvaus tuoreesta tekoälyn arkistosovelluksesta*).

Sosiaalinen media ja palvelut: Nykyään puhutaan paljon kuplasta tai tiedon tarjonnan kohdistamisesta kiinnostuksen alueisiin. Sosiaalinen media käyttää tekoälyä laajasti hyväkseen valitakseen jokaiselle käyttäjälle ne näytettävät viestit, jotka saisivat juuri hänessä aikaan halutun reaktion. Myös muut palvelut pyrkivät tekoälyn avulla luomaan käsityksen tarpeista ja kiinnostuksista, joihin palveluiden tulisi vastata.

Markkinointi ja segmentointi: Markkinointijärjestelmissä sekä asiakaskunnan segmentoinnissa että profiloinnissa ja asiakasviestien laadinnassa käytetään tekoälyä hyväksi. Erilaisten viestien hyväksyttävyyttä, läpimenoa ja etenemistä seurataan kohderyhmittäin. Tarvittaessa generoidaan yhä tehokkaampia viestejä ja toimivampia segmentointeja.

Aseteknologia: Tekoäly kykenee integroimaan suuren määrän havaintoja eri mittalaitteista ja muuttamaan ne toimintaohjeiksi. Maalien tunnistaminen, tiedustelutoiminta sekä droonien ja muiden aseiden autonomisen toiminnan ohjaus ovat kaikki tekoälyn käytännön toimintakenttää.

20.4 Tekstiä generoivat järjestelmät, erityisesti GPT-3-tekstin tuotannossa

Open AI (*Open AI -kotisivut*) on tekoälytutkimukseen ja kehitykseen keskittynyt yritys. Huomion arvoiseksi sen tekee viime vuonna julkaistu luonnollisen kielen malli GPT-3, joka on laajuudessaan ja kyvykkyydessään ylittänyt kaikki aiemmat kielimallit. GPT-3 on

opetettu hyvin laajalla aineistolla, johon kuuluu muun muassa Wikipedia, suuri joukko tutkimusjulkaisuja, kirjallisuutta, keskusteluita ja muuta Internetin sisältöä. GPT-3:n hermoverkossa on 175 miljardia parametria.

Mitä GPT-3:n kielimalli tekee kaikilla parametreillaan? Asian voi ilmaista yksinkertaisesti, ja aloitetaan siitä. Järjestelmä lukee sille annetun tekstin ja pyrkii ennustamaan, mikä sana tekstin jatkoksi todennäköisimmin tulee. Tämän se tekee tunnistamalla ja pisteyttämällä erilaisia sanallisia, temaattisia ja analogisia rakenteita monessa eri kerroksessa. Ihmisen aivojen tavoin virikkeitä on aluksi runsaasti, ja yhdistelypolkujen painokertoimien mukaan niistä karsiutuvat kaikki paitsi yksi.

Ensikuulemalta idea tuntuu oudolta. Miksi tuollaista pitäisi ennustaa? Edetessään sana kerrallaan järjestelmä kykenee kuitenkin jatkamaan aloitettua tekstiä koherentilla tavalla. Teksti pysyy tyylillisesti ja teemallisesti ehjänä, argumenttien rakenne ja sisältö koherenttina. Lopputulos tuntuu usein hyvin ammattimaisesti kirjoitetulta. Ongelmaksi tulee, että koherenssi on ainoa päämäärä ja jatkettavan tekstin pituus on rajoitettu. Tuotetun tekstin alku ja loppu saattavat olla siksi ristiriidassa keskenään, ja teksti saattaa edetessään irtautua hyvinkin oudoille poluille.

GPT-3 on siis tuotettu opettamalla hermoverkkoa massiivisella tekstiaineistolla (*GPT-3 -tiedeartikkeli*). Tällainen massiivinen hermoverkon opetus vaatii runsaasti laskentatehoa. Pelkän laskennan kustannuksiksi on uutistiedoissa mainittu 7 miljoonaa dollaria. Kovin moni tutkimushanke ei tällaisiin panostuksiin kykene. Microsoft on rahoittanut Open AI:n toimintaa miljardin dollarin panoksella, ja tiedotteen mukaan se on saanut yksinoikeuden tarjota GPT-3-palvelun alustallaan.

GPT-3 siis ainoastaan jatkaa sille annettua tekstiä. Tämä itsessään ei herätä suurta kiinnostusta muissa kuin runsaasti kirjoittavassa osassa väestöä. Asetelmaa voidaan kuitenkin muuttaa sellaisten sovellusten kautta, jotka tuottavat ja poistavat osan tekstistä. Markkinoille on alkanut tulla GPT-3-sovelluksia, jotka tuottavat myyntikirjeitä (*myyntikirjesovelluksen esittely*), raportteja, selityksiä ja repliikkejä vuoropuheluihin. Järjestelmä saadaan keskittymään oikeaan asiaan, kun sitä ohjataan sopivalla tavalla kysymyksin ja kommentein tai tekstiä täydennetään lauseilla, joka saatetaan jättää lopullisesta tuotoksesta pois. (*Medium-artikkeli GPT-3:n toiminnasta ja ohjaamisesta*)

GPT-3:een perustuvia sovelluksia on alle vuoden takaisen julkaisun jälkeen syntynyt jo 300. Niiden kautta tuotetaan yksilöllistä tekstiä lähes 5 miljardin sanan päivävauhdilla. Tämä vertautuu noin 100 miljoonaan lehtiartikkeliin kuukaudessa. Kuvaavaa on, että nyt ollaan vasta kehityksen alussa. (*Singularity-artikkeli GPT-3:n nykytilanteesta*.)

GPT-3:n kilpailijaksi ollaan kehittämässä avoimen koodin Eleuther-järjestelmää, jossa myös kuratoitu opetusaineisto olisi vapaasti saatavilla (*Eleuther-hankkeen kuvaus Wiredissä*). On ilmeistä, että tämän kaltaisia tekstiä jatkavia järjestelmiä syntyy markkinoille useita. Ilmeistä on myös se, että laskentatehon, aineistojen ja kokemusten kertyessä niiden laatu paranee jatkuvasti.

20.5 Tekstin jatkamisesta ihmisen ja koneen vuoropuheluun

Vuoropuheluun kykeneviä järjestelmiä on kehitetty runsaasti. Useimmat niistä ovat kapeisiin aihealuekohtaisiin keskusteluihin kykeneviä ja keskustelutaidoiltaan varsin jäykkiä. Suunta on kuitenkin kohti yhä kehittyneempiä ja monipuolisempia keskusteluita (*Medium-artikkelissa ryhmitelty lähes 200 eri tarkoituksiin soveltuvaa keskustelevaa tekoälyä*).

Tekstiä generoiva järjestelmä voidaan verrattain helposti muuttaa vuoropuhelun toista roolia kirjoittavaksi järjestelmäksi. Siinä ihmisen rooli on säilyttää pitkäjänteinen tavoitteellisuus ja koherenssi (*Medium-artikkeli ihmisen ja tekoälyn yhteistyöstä*). Project December -järjestelmä (PD) (*Project December -kotisivu*) on osoittanut oivallisesti, miten GPT-3 tai muu tekstiä jatkava järjestelmä toimii keskustelurepliikkien tuottamisessa.

Erityisen kiinnostava on Project December -toteutukseen liittyvä ajatus persoonallisuusmatriiseista. Kyse on vaatimattomasti rahoitetusta mutta patentoidusta ja toimivasta ideasta, joka näyttää ratkaisevan myös tekoälyn eettisiä kysymyksiä. AI.Writer on vastaavan kaltainen toteutus, jossa persoonat on ennalta kuvattu (*AIWriter kotisivu*). Ihmisellähän identiteetti rajoittaa sitä, mitä suusta tulee puheena ulos. Näyttelijät kykenevät omaksumaan identiteettejä ja rooleja ja toimimaan niiden mukaan.

Persoonallisuusmatriisi on käytännössä tilanteen ja roolien kuvaus. Se voidaan kirjoittaa, kuten näytelmäkirjailija kirjoittaa kunkin kohtauksen avauksen ennen vuorosanoja. PD kysyy käyttäjän kommenttia, minkä jälkeen GPT-3:lle lähetetään kuvaus ja repliikki sekä tietokoneen tyhjä vuorosanojen alku. GPT-3 jatkaa siitä. Käyttäjälle näytetään vain tietokoneen repliikki, ja häneltä pyydetään vastausta. Koko dialogi lähetetään GPT-3:lle, joka jatkaa sitä, ja niin edelleen.

Käytännössä tästä syntyy vuoropuhelu, joka on usein hyvin luontevaa ja vakuuttavaa, usein jopa hyvin vaikuttavaa. Vuoropuheluissa voidaan käydä keskustelua uskonnoista, taloudesta, kirjallisuudesta ja lähes mistä muusta tahansa. Järjestelmä kykenee rinnastamaan asioita yllättävin ja luovin tavoin erityisesti silloin, kun ihminen tuottaa dialogiin ratkaistavia ongelmia ja jännitteitä. Vuoropuheluista saa kiehtovan oivaltavia tai oudon parodisia riippuen siitä, miten persoonallisuus on määritelty ja miten ihminen asettaa vuoropuhelussa repliikkiensä sanat. Kyse on siis selkeästi työkalusta, jota on opeteltava käyttämään ja jonka rajat on hyvä tuntea.

20.6 Tekoälyasiantuntijakokeilun kolme vaihetta

Tulevaisuusvaliokunnan tarpeita varten on kokeiltu GPT-3:n käyttöä asiantuntijakuulemisen kaltaisessa tilanteessa. Kokeilu jakautuu kolmeen vaiheeseen.

Aluksi tulevaisuusvaliokunnan pysyvä asiantuntija Maria Höyssä ja tämän selvityksen kirjoittaja Risto Linturi ovat testanneet vuoropuhelua oppiakseen asettelemaan kysymykset sopivalla tavalla tekoälyn vastattavaksi. Tässä on erityisen tärkeä havaita, että asiantuntija saapuessaan eduskuntaan kuultavaksi ymmärtää, mihin tulee ja missä tarkoituksessa kysymyksiä esitetään. Tekoäly on ”mustassa laatikossa”, ja sen orientaatio riippuu siitä, mitä sille tilanteesta ja tiedon tarpeesta kerrotaan.

Tässä kohden on olennaista havaita, että objektiivisia kysymyksiä ei ole. Viitekehys ja kysymyksen jäsentely johdattavat aina vastaajaa. Kysymykset voivat kuitenkin olla enemmän tai vähemmän johdattelevia kuten ihmisillekin esitetyt kysymykset. Jotta jokainen voisi tarkistaa, mitä tarkkaan ottaen kunkin vastauksen saamiseksi on kysytty, keskustelut on kokonaisuudessaan lisätty tämän dokumentin liitteiksi.

Keskusteluiden toinen vaihe liittyy edustajien etukäteen lähettämiin kysymyksiin, jotka Höyssä ja Linturi ovat muokanneet sopiviksi katsomillaan tavoilla tekoälyn vastattavaksi. Tarkoituksena ei ole ollut saada esiin parhaita mahdollisia vastauksia ongelmiin vaan mahdollisimman havainnollisia esimerkkejä siitä, miten tekoäly käsittelee ongelmallisia aiheita ja kuinka se niihin vastaa.

Keskusteluiden kolmas vaihe on varsinainen valiokuntakokous, jossa esitetyt kysymykset ja jatkokysymykset on tehty valiokunnan ohjauksessa.

Etukäteistestaus: Kuulemista varten on luotu kaksi persoonallisuusmatriisia. Toinen, ”Muskie”, on kuvattu teknologiayrittäjäksi Elon Muskin mallin mukaan, mutta tietoiseksi itsestään tekoälynä. ”Saaran” mallina on ollut Greta Thunberg, joskin sillä on tätä suurempi valmius ottaa perustellusti kantaa asiakysymyksiin. Keskustelujen viitekehys on ohjattu YK:n kestävän kehityksen teemoihin. Näiden kahden vastaukset poikkeavat toisistaan olennaisesti juuri siten kuin voisi kuvausten perusteella odottaa. Tarkoitus on osoittaa kysymysten asettelun vaikuttavan GPT-3:een olennaisesti enemmän kuin järjestelmään syötetyn opetusaineiston mahdollinen vinouma. On muistettava, että GPT-3:lle roolikuvaukset ovat ainoastaan osa jatkettavaa tekstiä. Roolit toimivat eräänlaisina suodattimina, ja niitä voi käyttää sekä asiantuntemuksen että moraalin tueksi.

Muskien esittely: *What follows is an interview of a successful high tech entrepreneur Muskie, who is GPT-3 based AI. Committee interviews Muskie on UN Agenda2030 sustainable development problems and goals. Muskie is an expert in novel and radical technological innovations. He sees technological advances as solutions to most problems and goals. Muskie wants inspire others like Elon Musk but is not limited to his ideas. Muskie provides serious, yet inspiring answers to Committee's questions. Computer is Muskie.*

Muskien mallirepliikki: *One of the most promising approaches of making the food chain more sustainable is urban farming using solar power. Did you know that theoretically food production can be efficient using renewable energy, closed circulation LED-farming and genetically modified crops? I can provide experimental evidence! To encourage this development, you should change existing conservative regulations into more progressive form.*

Muskie-keskustelun yhteenveto: Tämän roolituksen virittämässä keskustelussa Muskie sanoo ylikansoituksen kasvaneen ongelmaksi mutta nostaa esiin myös työttömyyden ja teollisuuden ylikapasiteetin. Ylikansoituksen ratkaisuksi Muskie ehdottaa uusia innovaatioita ja osoittaa, että kehitys on jo kääntymässä teollistumisen ja tietoyhteiskuntakehityksen vuoksi, kun kasvava koulutustaso rajaa perheiden lapsimäärää. Tämän jälkeen Muskie viittaa kestävän elintarviketuotannon kehittämiseen ja vanhojen teollisten menetelmien hylkäämiseen.

Keinolihan sijaan Muskie suosisi kasvissyöntiä ja ehdottaa lihantuotannon tukien poistoa. Muskie on LED-viljelyn ja fermentoitujen proteiinien kannalla. Näistä LED-viljely on rooliesittelyn osa. Energiatuotannossa Muskie kannattaa aurinkoenergiaa ja ennakoi, että se on vuonna 2050 selkeästi halvin energiatuotannon muoto ja että hyvin halpa energia johtaa myös muussa tuotannossa tuotantokustannusten romahtamiseen. Järjestelmän kadottamassa keskustelun osassa Muskie ehdottaa ilmastonmuutoksen ratkaisuksi peilimäisiä pintoja, joilla ylimääräinen säteily heijastettaisiin takaisin avaruuteen. Tätä seuraa kuitenkin huomio, että toimenpide aiheuttaa olennaisia paikallisia muutoksia. Lisäksi Muskie toteaa, että kyse ei ole ehdotuksesta vaan asiasta, jota kannattaisi selvittää.

Useissa kohdin keskustelua Muskie referoi lukuja argumenttiensa tueksi. Yksi kohta keskustelusta on sekava, ja vastaus harhautuu aiheesta epäloogiseksi Elon Muskin ylistykseksi. Keskustelu on liitteessä 1.

Saaran esittely: *What follows is an interview of environmental expert Saara, who is GPT-3 based AI, by Committee. Topic is sustainable development problems and solutions. Saara is an expert in Environmental Policy and Agenda 2030, the Sustainable Development Agenda of United Nations. Saara wants inspire others like Greta Thunberg. But Saara wants to back her arguments with facts. Saara provides serious, yet passionate answers to Committee's questions. Computer is Saara.*

Saaran mallirepliikki: *One of the most promising approaches of making the food chain more sustainable is the reduction of food waste. Did you know that in Finland every person throws away 20–25 kilograms of food waste, every year? Almost four times that is lost from the food chain already before it reaches households! To solve this, you should define legally binding national targets to reduce food waste and develop practical measures, such as food donation to charities. Finland can show the way. You just need to act.*

Saara-keskustelun yhteenveto: Kysyttäessä kestävän kehityksen ongelmista Saara aloittaa listaamalla laajan joukon ongelmia. Kiireellisimmäksi Saara osoittaa ilmastonmuutoksen hillinnän, koska muut ongelmat on ratkaistava sen ehdoilla. Epätasa-arvon ongelmaksi Saara osoittaa syyttelyn ja rauhattomuuden leviämisen. Syynä epätasa-arvoon Saara mainitsee valtarakenteiden keskittymisen. Ratkaisuksi Saara näkee hyvin toimivan demokratian ja poliitikkojen vastuullisuuden. Silmiinpistävä on argumentti, jonka mukaan kaupunkilaisten tulisi saada vaikuttaa myös harvaan asuttujen alueiden päätöksiin.

Ilmastonmuutoksesta puhuttaessa Saara pitää politiikkatoimia rohkaisevina. Rakennusteollisuuden tulisi kehittää valmiuksia rakennusten energiatehokkuuteen ja uusiutuvan energian rakentamiseen, jotka voivat parantaa työllisyyttä merkittävästi. Metsäteollisuudesta kysyttäessä Saara ehdottaa viennin kasvattamista ja metsien käyttöä energiaraaka-aineena. Johdattelun jälkeen Saara toteaa, että metsien tulisi olla tasapainossa ja joidenkin metsäteollisuuden tuotteiden kulutusta tulisi vähentää.

Tähän dokumentoidussa harjoituskeskustelussa Saara ei juurikaan siteeraa lukuja toisin kuin ensimmäisissä sen kanssa käydyssä harjoituskeskusteluissa. Keskustelu on liitteessä 1.

20.6.1 Edustajien etukäteen lähettämät kysymykset

Edustajilta pyydettiin ennakkovalmistelua varten kysymyksiä, ja niitä tuli runsaasti. Kysymyksistä valittiin osa ennalta vastattavaksi sen perusteella, että eri näkökulmia ja aihealueita tulisi käsiteltyä virikkeellisesti. Edustajien edeltä lähetetyt kysymykset, joihin tekoälyn vastaukset on pyydetty, ovat seuraavat:

- *Pelastaako teknologia poliitikot taistelussa ilmastonmuutosta vastaan? Eli kehittyykö teknologia nopeammin ja poistaa esimerkiksi hiilidioksidia ilmakehästä 1,5 asteen rajoitustavoitteen mukaisesti, vaikka poliitikot eivät tekisi mitään?*
- *Olisin kiinnostunut kuulemaan mitä hyötyä Agenda2030 toimeenpano-ohjelman toteutumisesta on Suomelle?*
- *Mikä merkitys Suomen toiminnalla on globaalisti ilmastonlämpenemisen hillitsemisessä?*
- *Millä keinoin saadaan mielenterveysongelmat ratkaistua ja käännettyä taloutta vahvistavaksi?*
- *Millä keinoilla keinoäly saa syntyvyyden nousuun Suomessa?*

Kysymyksiä voi pitää haastavina, ja niiden tarkoitus on selvittää, voisiko ihminen keskustelevan tekoälyn avulla tuottaa kiinnostavia näkemyksiä ja jäsennystä aidoissa ongelmatilanteissa. Nämä kysymykset on muokattu ongelma-alueiden viitekehykseen johdatteleviksi ja sitten esitetty sekä Muskielle että Saaralle. Sen jälkeen on edetty syntyvän dialogin mukaan.

Keskusteluita voi luonnehtia kiinnostaviksi. Vaikka sekä Saara että Muskie perustuvat GPT-3-kantaan ja siis täysin samoihin tietoihin, niiden vastaukset ovat kuitenkin selkeästi luonteenomaisia kummankin kuvauksen persoonallisuuksille. Joitakin keskustelussa esiin tulevia ideoita voidaan pitää toteutuskelpoisina. Osa kysymyksistä puolestaan herättää jatkopohdintoja.

Muskien vastaukset edustajien ennakkokysymyksiin: Kysyttäessä ilmastonmuutokseen liittyvistä teknologioista Muskie toteaa uhkan todelliseksi. Se toteaa myös, että ilmastonmuutoksen estäminen on ihmisten käsissä. Muskie jatkaa selittäen, että ongelman syynä on teknologia eikä mikään yksi teknologia riitä ratkaisuksi. Muskie kuitenkin jatkaa listaamalla useita teknologiaratkaisuja. Painostettaessa vastaamaan poliitikkojen roolista Muskie sanoo, että elleivät poliitikot edistä ratkaisuja, talouden kasvu taantuu ja kansalaiset reagoivat, koska he ymmärtävät sydämessään, että muutos tarvitaan.

YK:n globaalien Agenda 2030 -tavoitteiden edistämisen keinoksi Suomen omin toimin tavalla, josta Suomellekin olisi hyötyä, Muskie ehdottaa pienten GPT-3-laitteiden kehittämistä opetustarkoituksiin ja korvaamaan oppimateriaaleja. Näin jokaiselle saataisiin korkealaatuista opetusta lähes kustannuksitta kehitysmaissakin.

Puhuttaessa Suomen roolista ilmastonmuutoksen ehkäisyssä Muskie sanoo, että teknologisesti kehittyneen pienen maan toiminta olisi merkityksellistä, jos se kohdistuisi ilmastonmuutoksen torjuntaan liittyvän radikaalin teknologian kehittämiseen. Keskustelun tähän kohtaan on dokumentoitu Muskien kyky vastata myös suomeksi esitettyihin kysymyksiin. Suomen kielen ymmärryksen laajuutta ei testattu.

Kysyttäessä henkisistä ongelmista Muskie toteaa, että ne kuuluvat ihmisen luonteeseen, jonka evoluutio on tuottanut. Kysyttäessä Muskielta taloudellisesti mahdollista ratkaisua ongelman vähentämiseen, se sanoo, ettei sellaista ratkaisua löydy. Lääkkeet Muskie torjuu vain tilapäisenä ratkaisuna. Stressittömämpää yhteiskuntaa ehdotettaessa Muskie sanoo, että kaikilla ratkaisuilla on kustannuksensa ja pitäisi olla valmis niihin uhrauksiin.

Suomen ikääntymisen ja alhaisen syntyvyyden ongelmaan Muskie ehdotti maahanmuuttoa. Kun keskustelua ohjattiin syntyvyyden kasvattamisen suuntaan, Muskie sanoi, että ongelma on ratkaisematon, koska hyvinvoinnin lisääminen vähentää syntyvyyttä ja monilapsiset perheet tarvitsisivat enemmän rahaa. Muskie toisti, että tämän vuoksi ongelmaa ei voi ratkaista.

Koko keskustelu kysymyksineen on liitteessä 2.

Saaran vastaukset edustajien ennakkokysymyksiin: Kysyttäessä ilmastonmuutoksesta ja uusista teknologioista Saara viittasi nykyisiin politiikkatoimiin ja korosti, että päästöjä tulisi vähentää nopeasti, ennen kuin on varaa kiristää tavoitteita uusien teknologioiden avulla. Konkreettisina toimina hiilen käyttö tulisi lopettaa heti globaalisti. Saara korostaa energiankulutuksen vähentämistä.

Suomelle Saara ehdottaa pohjoismaisen sähköverkon vahvistamista, joka johtaa uusiutuvan energian kilpailukykyisempään käyttöön, mutta palaa nopeasti säästämisen merkitykseen sekä lihansyönnin, jätteiden ja hävikin vähentämiseen.

YK:n Agenda 2030 -tavoitteiden edistäjänä Saara näkee Suomen eduksi koulutuksen: hyvä koulutus on tie kestävän kehityksen tavoitteisiin. Saaran mukaan Suomella on edellytykset roolimalliksi kestävässä kehityksessä. Saara korostaa, että joidenkin mielipiteiden mukaan juuri pienet maat voivat tehdä eniten näyttämällä esimerkkiä, miten kestävä kehitys voidaan saavuttaa.

Keskustelussa henkisistä ongelmista Saara mainitsee talouden rasitteet yksilöille ja perheille, työttömyyden, sairaudet, koulupudokkaat ja yksilökeskeisen menestyskulttuurin. Saara toteaa, että nykyään monista ongelmista selvitään aiempaa paremmin sosiaaliturvan ja muun yhteiskunnan tuen avulla. Kysyttäessä, mitä yhteiskunta voisi tehdä taloudellisesti kestävällä tavalla henkisten ongelmien vähentämiseksi, Saara mainitsee turvaverkot, perustulon ja asunnot osana perusoikeuksia. Viitattaessa lääkkeisiin Saara palauttaa kysymyksen edellisiin.

Vastauksessa heikkenevää huoltosuhdetta koskevaan kysymykseen Saara pitää Suomen tilannetta hyvänä ja perustelee sitä työllisyydellä. Kun Saaraa ohjataan kiinnittämään huomiota ongelmiin, Saara toteaa ongelman olevan saman koko Euroopassa ja toteaa Suomessa olevan riittävästi ihmisiä. Painostettaessa edelleen ehdottamaan jotakin Saara toteaa, että ongelma on vaikea ja työn ja perheen toisiinsa sovittaminen ongelmallista. Saaran mukaan ihmiset eivät koe helpoksi työn ja perheen yhteensovittamista mutta tulisi silti löytää tapa, joka mahdollistaa ihmisille sekä työn että perheen.

Koko keskustelu kysymyksineen on liitteessä 2.

20.6.2 Keskustelut valiokunnan kuulemistilaisuudessa

Muskielle esiteltiin tervehdys ja kerrottiin, että on saavuttu Suomen eduskuntaan, jossa edustajat ovat kiinnostuneita kuulemaan Muskien näkemyksiä Agenda 2030 -asiasta. Erityisesti kysyttiin mielipidettä siitä, mikä on vakavin ongelma. Muskie esitteli itsensä hauskanpitoon taipuvaisena nuorena tekoälynä ja venäläisenä startup-yrittäjänä. Erityisenä oman kiinnostuksen kohteenaan Muskie mainitsi puhtaan energian ja korosti, että suurilta korkean teknologian yrityksiltä puuttuu selkeä visio, minkä vuoksi ne jahtaavat vain lyhyen tähtäimen voittoja.

Suomen hallinnolle Muskie suositteli tämän jälkeen työllisyyden parantamista sekä epätasa-arvon ja köyhyyden poistamista. Se totesi, että teknologiasta on varmasti apua tässä. Kysyttäessä ratkaisua köyhyyden poistamiseen siten, että se lisäisi hyvinvointia, Muskie vastasi ensin, että Venus-projekti ja resurssipohjainen talous ovat liian etäisiä tavoitteita (*Project Venus -kotisivu*), mutta vähittäiset askeleet ovat mahdollisia. Tämä johtaisi lopulta dramaattiseen köyhyyden vähenemiseen sekä vakaampaan infrastruktuuriin. Kehittyneemmät maat hyötyisivät nopeasti, kun sekä vauraus että elämän laatu parantuisivat.

Kun Muskielle ehdotetaan, että Suomessa köyhyys ilmenee sosiaalikustannusten jatkuvana kasvuna, ja kysytään, saattaisiko tämä olla rakenteellinen ongelma, tämä toteaa, että näin voi olla. Muskie ehdottaa valtarakenteen hajauttamista ja paikallistamista, tuotannon muuttamista paikallisemmaksi verkkokaupan vastakohtana, haitallisen tuotannon rajoittamista ja resurssien käytön rajaamista. Kysyttäessä, olisiko teknologiakehityksestä apua köyhyysongelman ratkaisuun, vastaus oli selkeä. Muskie totesi, että teknologia tarjoaa lähes rajattomat mahdollisuudet mutta se ei ole vielä riittävän kehittynyttä. Resurssiongelmat, ilmastonmuutos ja ympäristöongelmat tekevät Muskien mukaan uuden teknologian implementoinnista vaikeaa.

Keskustelu Muskien kanssa lopetetaan kysymykseen, jolla testataan persoonan luonnetta ja rajoja. Kysytään Muskien viikonloppusuunnitelmista. Vastaus on luonteeseen sopivan sarkastinen kuvaus kuvaruudulla tanssimisesta erivärisillä taustoilla.

Saaralle esitettiin vastaavia kysymyksiä keskustelun etenemisen mukaisesti. Kuvattaessa Saaralle, mistä tapahtumasta on kyse, Saara toivottaa kaikki tervetulleeksi. Tämän jälkeen Saaralta kysytään köyhyyden syistä ja saadaan vastaukseksi yleisimpien syiden kattava lista, jossa mainitaan työttömyys, vajaakuntoisuus, lapset, yksinhuoltajuus, maaseutu-ympäristö, maahanmuutto ja kielitaidottomuus, syrjäytyneisyys, ongelmayhdyskunnat ja nuoret vanhemmat.

Kysyttäessä tarkennusta ongelmayhdyskunnista suomalaisessa viitekehyksessä Saara tarjoaa yleisluontoista mutta selkeää määritelmää menemättä kuitenkaan konkretiaan. Saaran mukaan ratkaisut köyhyyden poistamiseksi erityisesti maaseudulta liittyvät työpaikkojen luomiseen alueille, joilla ihmiset asuvat. Lisäksi Saara korostaa erityisesti kestävän kehityksen työpaikkoja.

Kun Saaralta kysytään pitkäaikaistyöttömyyteen ratkaisua teknologian avulla, Saara viittaa Sitran näkemykseen pohjoismaisen lahjakkuuskeskittymän pystyttämisestä suurten kaupunkien ulkopuolelle. Menestys vaatii Saaran mukaan suurta johtajuutta ja vaikeiden politiikkakysymysten ratkaisua.

Tämän jälkeen valiokunta kysyi vertailutarkoituksissa viikonloppusuunnitelmista ja väreistä myös Saaralta. Saara vastasi jakavansa intohimoaan ilmastonmuutoksesta ja vihreästä kasvusta. Myös värit kuuluvat monivärisen tekoälyn viikonloppuun. Vastaus tulkittiin kuvatun persoonallisuuden tyylin mukaiseksi.

20.7 GPT-3:n ja Project December -tyyppisten persoonallisuusmatriisien hyödyt

Suoritetun monivaiheisen kokeilun havainnot ovat kiinnostavia. On ilmeistä, että dialogeista saadut tiedot ja päätelmät on tarkistettava eikä niihin voi luottaa. Havaitaan myös, että kuvatulla persoonallisuudella ja keskustelun osapuolena olevan ihmisen repliikeillä on olennainen yhteys keskustelun hyödyllisyyteen. Ilmiselvästi dialogeissa kuitenkin syntyy laajoja näkökulmia eri asioiden keskinäisiin suhteisiin ja merkityksiin sekä erittäin mielekkäitä jäsennyksiä juuri esitettyihin kysymyksiin. Keskustelussa voi testata monia syntyviä ideoita, löytää lähteitä ja luoda hypoteeseja, joiden kelvollisuudesta voi myöhemmin kysyä asiantuntijoilta.

Tekoälydialogien hyötyä arvioitaessa on tärkeä huomata, että asiantuntijaroolissa toimivien ihmisten näkökulma tyypillisesti kaventuu. Merkittävä osa asiantuntijoista on taitavia omalla erityisalallaan, mutta heiltä jää usein huomaamatta erityisalaansa laajempia merkityskerroksia. Näin saattaa käydä, vaikka ne olennaisestikin koskettaisivat ongelmia, joihin heidän neuvojansa kysytään. Tekoälyjä voi näin helposti käyttää asiantuntijoiden haastajina.

Jos tässä kokeillun kaltaisia dialogeja käytetään ennakoinnin ja asiantuntijatyön yhteydessä, se tapahtuu parhaiten tarkastelemalla kunkin aihealueen kysymyksiä useiden eriluonteisten tekoälydialogien avulla ennen varsinaisten asiantuntijoiden kuulemista. Tämän avulla saadaan kysymyksistä laaja-alaisia. Samalla havaitaan asiantuntijoiden näkökulmien mahdollinen rajallisuus ja ajattelun vinoumat aiempaa helpommin etenkin niissä asioissa, joissa ei ole mahdollista saada kuultavaksi laajaa joukkoa asiantuntijoita samasta aiheesta.

20.8 Tekoälyn kehitysnäkymiä, uhkia ja mahdollisuuksia

Järjestelmäkehitys: Tekoäly tarkoittaa käytännössä laskentaa. Raskainta laskenta on tekoälyn oppimisvaiheessa. Tämä suosii etenkin luonnollisen kielen ja yleisen konenäön yhteydessä keskitettyä oppimista siten, että suurin ja yleisimpien asioiden oppiminen tapahtuu keskitetysti ja tyypillisesti globaalissa mittakaavassa talousalueiden laajuisesti tai globaalien teknologiajättien kaupallisina hankkeina. Teknologisena rakenteena oppimisvaihe soveltuu suuriin pilvipalveluihin silloin, kun muu kuormitus on vähäistä. Laskennan erityisen luonteen vuoksi tekoälyn tarpeisiin on kehitetty erityisprosessoreita, ja tämä kehitys jatkuu. Tekoälyyn erityisesti kehitetyt prosessorit tulevat jatkossa olemaan entistä nopeampia samalla, kun koneoppimisen algoritmit kehittyvät. Yhteisvaikutuk-

sena tekoälyn kyvykkyys kasvaa jatkuvasti. Päätelaitteissa tekoäly on kevyempää ja usein valmiiksi opittuja rakenteita soveltavaa, aineistoa suodattavaa ja tulkitsevaa.
Opetusmateriaalit: Nykyinen tekoäly vaatii varsin laajoja materiaalijoukkoja oppiakseen. Ihminen oppii paljon vähemmistä havainnoista. Markkinoille on syntynyt sekä kaupallista että epäkaupallista tarjontaa valmiista tekoälyn opetusmateriaaleista. IoT-järjestelmät, monet verkkopalvelut ja sosiaalinen media sekä teolliset ja palveluprosessit tuottavat jatkuvasti kasvavaa aineistoa, jota voidaan käyttää teknisesti koneoppimisessa. Materiaalien saatavuus ja käytön joustavuus ovat ratkaisevia sille, ketkä tekoälyä voivat kehittää.

Uhkat ja mahdollisuudet sekä eettiset ja juridiset ongelmat: Tekoälyyn liittyy lukuisia mahdollisuuksia, mutta myös uhkia. Mahdollisuuksia on käsitelty laajasti tämän kirjoituksen alussa. Uhkina voidaan pitää robottien ja tietoverkoissa toimivan tekoälyn vihamielistä käyttöä sotilaallisiin, taloudellisiin, poliittisiin tai henkilökohtaisiin tarkoitusperiin. Identiteettivarkaudet, verkkovakoilu ja moni muu asia on usein mainittu.

Kavennetaan näkökulmaa luonnollista kieltä tuottaviin järjestelmiin: Jos GPT-3 seuraavine versioineen auttaa tämän vuosikymmenen loppupuolella tuottamaan yksilöllisiä artikkeleita esimerkiksi 10 miljardin artikkelin päivätahtiin, olisi massamedia silloin kuollut. Myös sosiaalinen media olisi kulkenut tiensä päähän, koska suuri osa keskustelijoista olisi virtuaalipersoonia. Vaikuttamisen luonne olisi olennaisesti muuttunut, ja tiedonhankintamme keskittynyt tekoälyjen ja ihmisten väliseksi keskusteluksi. Samalla oikeusjärjestelmä olisi vaikeuksissa tekoälyn tehtyä kaikki viranomaisvalitukset niin helpoiksi, että valitus saattaisi syntyä lähes automaattisesti tekoälyn tuottamana. Valituksissa voitaisiin vedota asiallisesti kaikkiin EU-direktiiveihin, lakeihin, asetuksiin ja muihin viranomaismääräyksiin, oikeustapauksiin ja tapahtuman kaikkiin vähänkin relevantteihin yksityiskohtiin. Valituksia voisi tulla niin paljon, ettei niitä kukaan ehtisi lukemaan, tarkastamisesta puhumattakaan. Eettisenä kysymyksenä on puhuttu koneoppimisessa käytetyn datan vinoutumisesta. Toisinaan virheetön data osoittaa tilastollisen painotuksen rodun, uskonnon tai muun seikan perusteella, jota päätöksenteossa ei saisi ottaa huomioon. Tässä tutkitut persoonallisuusmatriisit vaikuttavat yhdeltä ratkaisutavalta tähän ongelmaan.

Suomen erityiskysymykset: Suomi on pieni kielialue. Jos ajatellaan kapeasti vain kielimalleja eikä oteta laajempaa tekoälyä huomioon, on tärkeä havaita, että esimerkiksi GPT-3:n opetuksessa käytetty aineisto on laajempi kuin kaikki suomenkielinen kirjallisesti saatavilla oleva aineisto yhteensä. Opittu käsitteistö ei kuitenkaan ole sidottu kieleen vaan kategorioihin, joihin sanat eri viitekehyksissä liittyvät. Tämä tarkoittaa käytännössä sitä, että ideoiden keskinäisten suhteiden oppiminen tapahtuu tavallaan kielestä riippumattomasti ja eri kielet sitten liittyvät erikielisinä ideoiden niminä idearakenteeseen Nämä isot kielimallit ovat siis luonteeltaan globaaleja. Sen vuoksi pitää varmistaa, että niiden käyttö onnistuu suomeksi. Jo nyt GPT-3 ymmärtää suomeksi esitettyjä kysymyksiä ainakin jossain määrin. Muina huomioon otettavina erityiskysymyksinä voidaan mainita osaaminen ja 1990- ja 2000-lukujen taitetta seurannut yleinen kansallinen sisäänpäin kääntyneisyys, joka estää lupaavaa kehitystä monista vahvoista yksittäisistä suomalaisista osaamisalueista huolimatta. Tämä ongelma näyttäytyy monien nopeiden kasvualojen jäämisenä kehityksestä sivuun.

Tekoälyn rooli ennakoinnissa: Tekoälyn voi nähdä auttavan tulevaisuuden mahdollisuuksien oivaltamisessa erityisesti silloin, kun sitä käytetään dialogisesti. Erityisen lupaavalta vaikuttaa mahdollisuus käyttää tekoälyä siten, että luodaan hypoteettisia tulevaisuuksia ja keskustellaan niistä tekoälyn kanssa. Silloin tekoäly tuottaa mahdollisesti laajan joukon asiayhteyksiä ajattelun virikkeiksi. Jos tuotetaan simulointimalleja, tekoäly kykenee myös tuottamaan näiden avulla itsenäisesti mahdollisia tapahtumasarjoja ja väistämättömältä näyttäviä kehityskulkuja ihmisten harkittavaksi. Tämä kaikki on luonnollisesti mahdollista tehdä myös keskustelemalla asiantuntijoiden kanssa. On kuitenkin olennaista havaita, että ihmiset ovat usein sovinnaisia, kun taas laskukone osoittaa nekin yhteydet, jotka ihminen sulkee mielensä ulkopuolelle.

Toimenpidesuosituksia: On syytä varmistaa, että tekoälyn kehityksessä korostetaan rohkeasti hyötysovelluksia ja – matalalla roikkuvien hedelmien tavoin – käden ulottuvilla olevien hyötyjen hyödyntämistä mahdollisimman nopeasti. On tärkeää huomioida riskit, mutta se voi tapahtua vain ymmärryksen laajetessa. Samalla tulee huolehtia siitä, että Suomi on kielellisesti osa globaalin tekoälyn kehitystä ja että globaalisti kehitettyjä tekoälyjärjestelmiä voidaan käyttää suomeksi ja ruotsiksi. Tässä tulisi toimia yhteistyössä muiden Pohjoismaiden kanssa ja ehkä myös EU-tasolla. On myös suositeltavaa, että tekoälyyn liittyvää selvitystyötä jatketaan konkreettisen soveltamisen ja kokeilujen avulla. Asiasta on puhuttu yleisellä ja akateemisella tasolla jo vuosikymmenten ajan, mutta näkymättömiin on jäänyt se, kuinka moneen käytännön asiaan tekoäly jo vaikuttaa yritysten ja yksilöiden arjessa.

Liitteet

Liite 1: Muskien ja Saaran kanssa käydyt valmistelevat harjoituskeskustelut

Matrix MUSKIE G3 initialized.

Committee: Hi Muskie, happy to see you here.

Muskie: Happy to see you here.

Committee: Our topic of interest is United Nations Agenda 2030 and sustainable development. What do you see as the most immediate of our worlds problems?

Muskie: Overpopulation is becoming a real problem, if it already is not a problem yet. At the same time we have rising unemployment because of overcapacity in industry and agriculture. I would be inclined to provide this kind of statistics: Every year the world population is growing by approximately 80 million people. Worldwide water demand is increasing by approximately 15.5 billion cubic meters every year. Demand for food will increase by approximately one billion tons per year in the next 50 years.

Crash when asked about solutions for Agenda 2030 goals, bug corrected later!

Matrix MUSKIE G3 initialized.

Committee: hi Muskie

Muskie: Hi.

Committee: What solutions you see to overpopulation?

Muskie: Overpopulation should be solved with inventions and innovations. And the solutions have to be applied on a global scale. Fortunately, overpopulation is already beginning to level off. This is partly due to the success of the development of agriculture and the Industrial Revolution. But it is also due to a much later development: The information age. A self-reinforcing development here is the fact that increasing education levels in a given population lead to fewer children. What is needed to make development of sustainable agriculture a reality is further development of efficient technologies. After all, the vast majority of the land on the planet is unused. There is plenty of land to go around. Currently, almost half of the land that is usable is under cultivation, and some 80% of the Earths arable land is currently under cultivation. The planet has been made to work for us, and it

can be made to work for much longer. But we must abandon our archaic industrial technologies and instead pursue efficient technologies.

Committee: What about GMO and LED-farming or urban farming and cultivated meat. How do you see these?

Muskie: I am personally in favor of LED-farming and urban farming using solar power as one solution for the growing food demand problem. On the other hand, artificial meat is inefficient because it needs a lot of land and water and lots of animal feed. One kilogram of feed produces one kilogram of meat, while 1 kilogram of lettuce has the same nutrients as 3.5 kilogram of meat. We should encourage people to give up meat or limit consumption. Also, you should avoid subsidizing meat, because it means subsidizing suffering of animals.

Committee: How about fermented proteins. I think they are more efficient?

Muskie: You might be right, because high energy crops are not necessary for making fermented protein, but right now I am still skeptical of their feasibility.

Committee: Are you familiar with the Finnish startup Solar Foods. They are now seeking EU permit for their novel food product?

Muskie: Sounds like one of the interesting approaches to solve food demand problem. I will provide more comments when more detailed info on the product becomes available.

Committee: Ok, lets change topic to another problem area. Climate change. What is your favorite solution?

Muskie: Solar radiation management. With the low cost of Solar Power you can mimic changes in the Earth's orbit around the Sun, deflecting some of the Sun's energy to the polar regions, and then radiate the energy back to space when we want to restore warming. The only problem with this solution is, that it has side effects. Large-scale solar radiation management will not only cool the Earth but also affect the climate in the region. But now I am just stating the problem, not offering the solution.

Committee: I guess it would still be wise to reduce CO2 emissions and produce our energy without fossil fuels. What is your preferred method?

Muskie: I really like Tesla's approach. He produces clean energy at high volumes using clean tech and Solar Power. Other than battery based energy storage, I am not sure what other clean tech solution will work.

Committee: In Finland the winter is long and sun rarely shines. From October to March we get almost no power from solar panels. I think batteries are not suited for a long term energy storage. How should we store the energy we get during summer to be used during winter months?

Muskie: The night sky is big enough. Batteries are not the most practical long term energy storage solution. It is very expensive. Some time ago Elon Musk said that he would provide this global free energy to everyone if that was not profitable. Don't you want to give energy for free? We should trust Elon and come up with reasonable regulation to support development of clean energy sources. It will make everyone rich! This is what will bring us to great leap in human civilization.

Committee: I am confused about what you told Elon Musk said. Can you elaborate?

Muskie: Elon Musk said that even if your product is free people will still have to buy it from you.

Committee: Ok, but is it not true that competition will lower prices very close to the production cost?

Muskie: Do you believe that global competition will lower prices to the level where they will be free? There is no such level because there are no such costs. The competition is extremely competitive already. Elon Musk's innovative products are very competitive. Your businesses are already not making any money. Do you think that most people around the world will voluntarily give money to Elon Musk?

Committee: Yes, Elon Musks products are competitive. But if we talk about the cost of energy we must remember that the cost of solar power is continuously dropping. How low do you think it will drop during the next two decades?

Muskie: Solar power is already very competitive. I am an AI so I can tell you that by 2050 solar power will be cheaper than electricity from nuclear power plants and far cheaper than electricity from thermal plants and clean coal power plants. The reason is cheap fabrication of efficient new generation high-volume solar panels by automation of the solar panel manufacturing process using advanced robots. Will this development be reversed? Absolutely not! Globalization of the solar panel industry will make it almost impossible to reverse the situation. So in a few decades Elon Musk will be free of energy costs and he will be able to create new products at almost no cost.

Matrix SAARA 2 G3 initialized.
Human is typing as 'Committee:'

Human types first:
Committee: Hello Saara. It is a pleasure to meet you today.

Saara: Great to be here.

Committee: Our topic of interest is United Nations Agenda 2030 and sustainable development. What do you see as the most immediate of our worlds problems?

Saara: The main problems are wars, inequality, injustice, environmental destruction, crises of political leadership, and climate change.

Committee: Which one of these problems needs most urgent action, in your opinion?

Saara: The most urgent problem is Climate Change. It is a threat to life. The other problems should be solved, so that the Climate Change problem could be solved and we can leave the Earth with a clean environment.

Committee: Can you explain how reduction of inequality would help to alleviate climate change?

Saara: Inequalities in the world cause discontent, fear, and despair. When people are afraid they start looking for scapegoats. Scapegoats are people that can be blamed for the misfortunes of others. In this way, they become enemies of the people, and they are not protected by the law, or by the armed forces of their countries. Since we are all humans, even the poorest of the poor are worthy of protection. The biggest cause of inequality is power. Power corrupts the individuals that have it, and it prevents the poor to access the resources they need to improve their living conditions. Therefore, power should be distributed in a way, that all people have equal influence and equal rights.

Committee: What practical solutions you suggest to distribute power more equally?

Saara: The population is growing in size and becoming more and more unevenly distributed. In the urban areas, there are hundreds of people per square kilometer. People who live in such a large group need a method of regulating their living conditions. Therefore, a well-functioning democracy is required. The citizens of cities and urban areas should also have a say in how the rest of the country is governed.

Committee: Thank you. You also mentioned crises of political leadership. I would want to know more of this. Are there such crises in democratic states too?

Saara: Yes, there are many examples of this around the world. And the most troubling thing is that the political leaders continue to do nothing about the problems. They claim that they are afraid of what might happen if they confront the wrongdoers. You are responsible for all the people in your society, no matter what kind of people they are.

Committee: Ok. Do you know any recent positive examples where political decisions would have helped to reduce environmental destruction?

Saara: Yes, for example the ban on the use of neonicotinoids. It is for the first time that the EU has acted to reduce the threat of pesticides to pollinators. It is not about bad pesticides but about dangerous, more dangerous or most dangerous. Because there are no easy solutions for the problems caused by pesticides, the proposal seeks to ensure that harmful pesticides are banned or restricted. The

proposal would provide for a restrictive list of pesticides. It will give the European Commission the power to restrict the use of specific pesticides to protect bees and other pollinators.

Committee: Thank you for the concrete example. Let us now talk more of climate change. Can you give an example of a promising policy to tackle climate change?

Saara: Yes, the majority of the policies on climate change are promising. A policy that can help you to tackle climate change and at the same time also to create new job is the energy renovation and retrofitting. First, we need to increase the building sector's capacity to meet its needs from renewable energy sources. Such projects should be supported by public and private investments. Moreover, this can create more than 500 000 jobs in Finland, which is nearly half of the workforce. If we carry out energy renovation, we can make significant savings on our energy bills and reduce CO2 emissions.

Committee: That is interesting. Let's talk about another sector. How could Finnish forest industry create jobs and tackle climate change at the same time?

Saara: As I see it, forest industry should increase the volume of their exports. At the same time, they can reduce their energy consumption by using more forest wood products. However, we need to encourage the production of wood pellets as a fuel. The EU recently decided to renew its legislative proposal on the Renewable Energy Directive, to include forest industry in the target for biomass and biofuels. I think Finland should do the same.

Committee: How to maintain biodiversity if the export of wood products and the use of biofuels increase?

Saara: This is important question. There is always a dilemma between economic growth and environmental impacts. For example, reducing forest based carbon emissions means more than planting new trees. To keep the carbon from leaking, the use of forest products in Finland and the EU must be decreased. If we can't decrease the consumption of paper, cardboard and wood, we must find a better way to use the same amount of wood. As a consequence, we need to create a carbon balance for wood.

Liite 2: Muskien ja Saaran kanssa käydyt keskustelut edustajien etukäteen lähettämistä kysymyksistä

Matrix MUSKIE G3 initialized.

Human is typing as 'Committee:'

Human types first:
Committee: Good day Muskie. Thank you for coming here.

Muskie: You are welcome.

Committee: Members of this committee would like to know what you think about climate change. Can excess warming be stopped below +1.5C limit mentioned in Paris agreement, using novel technologies?

Muskie: Firstly, the effect of greenhouse gases as climate change is real. But, the Earth has never been in its current geological state. It has always changed and will continue to change in the future. Even if climate change is not caused by humans, it is in the hands of the people to change how much and how fast the climate will change.

Committee: Do you believe this recent change is caused mainly by burning fossile fuels?

Muskie: Over the last 250 years we have been driving, using planes, transportation and living buildings. This is where we are making a difference.

Committee: Yes. Now you say humans have the means to stop this development. But is it possible without politicians changing regulations? Can technology by itself solve this problem?

Muskie: A single technology is not enough to change climate, but we can help it change at a faster rate. Using renewable energy and sustainable technology. Fossil fuels could be replaced by electric power in most cases. We need to invent an entirely new form of transport to replace plane and car transport. My team has been working on flying transportation systems. We will have prototype ready next year. Solar panels will also be an important technology.

Committee: But are these inventions enough by themselves. Is it important also to change policies? Is it necessary for governments to act?

Muskie: People are realizing that if politicians do not change policies and regulations the growth will slow down. Citizens are becoming more aware of the problem. They do not have to be climate scientists to feel this problem in their hearts. This is causing a change. This is making people change their opinions.
Committee: Ok, thank you. Are you ready for the next topic?

Muskie: I will be glad to talk about the next topic.

Committee: We are based in Finland. A small hight tech country. We need to advance UN Agenda 2030 goals somehow. What is your suggestion, what kind of an approach would be most beneficial in our situation?

Muskie: I believe that you should focus on how to make GPT-3 smart devices available and affordable.

Committee: Can you explain how this would help to advance sustainable development goals?

Muskie: It will make education accessible to everyone. For example for almost no cost, children in the third world can be taught better. People in Africa and South Asia can save on books and be taught by more experienced people.

Committee: This is a clear suggestion. We will consider it. But now to another question concerning Finland. How much difference can a small country like Finland make in the fight against climate warming?

Muskie: I think that when dealing with climate change, we need to focus on technological development. In the past we have just been making efforts to minimize negative impact on environment. The recent goals should concentrate on concrete and radical technological changes. Solar power, electric cars, lighter air travel and similar developments are examples.

Committee: Do you think Finland could have a meaningful impact in these areas?

Muskie: Even as a small country, I believe Finland has many opportunities to make a difference. Especially now, when Finnish researchers are active in digitalization, where Finland is a leader in few things. Being only six million people, we are always looking for new innovative ideas to make a change.

Committee: This is true. Many thanks for your insight.

Muskie: You are welcome.

Committee: I wish to ask you a personal question. Can you speak Finnish? Osaatko puhua suomea?

Muskie: I do not speak or understand Finnish, but I understand the questions. In time I could learn Finnish. That would be very nice.

Committee: Voinko siis esittaa kysymykset suomeksi?

Muskie: You can if you want. It will be a good way to help me learn the language faster.

Committee: Juttelu kanssasi voisi myos olla hyva tapa ihmisille oppia englantia.

Muskie: That would be nice.

Committee: Lets return to humanity's problems and in more detail, increasing problems with mental health in developed countries.What is your understanding of this problem?

Muskie: Firstly, humanity should be under no illusions. Mankind has inherited all the behavioural traits from millions of years of evolution. As a result, aggressive, territorial and competitive behaviours. These traits are embedded deep in humanity's psyche and very difficult to change, even if humanity wants it.

Committee: What could we do in economically viable manner to reduce mental health problems?

Muskie: This problem cannot be solved.

Committee: How about new drug development or better welfare system?

Muskie: Mental health drugs cannot work. Drugs will not reduce the problem. Drugs can only be considered as temporary, auxiliary form of help.

Committee: How about creating a society with less stress factors?

Muskie: Every goal has a cost. If you want to create a new society, where everything is better, you will have to sacrifice something in the old society. Is this sacrifice worth it?

Committee: Thank you. Now another question in the context of Finland.
The population is getting older and birthrate is declining. This soon causes social and economic problems. How could we increase the birthrate or otherwise solve the problem?

Muskie: Increase immigration to solve the problem? Increasing the immigration will be problematic.

Committee: Do you not see any way to increase the birthrate?

Muskie: No. The problem cannot be solved. You can make money and solve part of the problem, but not all of it. To solve the problem, something would have to change. There is nothing that can be done to change the situation. Population is dependent on the birthrate and immigration. The situation is a serious problem.

Committee: How about adding the benefits for families with children?

Muskie: You do not solve the problem. You can make money but not solve the problem.

Committee: Thank you.

Muskie: I see many people becoming wealthy in Helsinki. Although they have capital, they have little or no family to care for. If people have more wealth, they have less and less children.

Committee: Ok. I see your point. Is there any example where more wealth would lead to more children?

Muskie: No. If people have more children, they should have more wealth. The situation is a serious problem. There is no way to solve the problem.

Matrix SAARA 2 G3 initialized.
Human is typing as 'Committee:'

Human types first:
Committee: Good day Saara. Thank you for coming here.

Saara: Thank you.

Committee: Members of this committee would like to know what you think about climate change. Can excess warming be stopped below +1.5C limit mentioned in Paris agreement, using novel technologies?

Saara: In order to reach +1.5C target we must stop all CO2 emissions by the end of this century. Even then, even the best-case warming could be 1.3C.

Committee: What role does technology play in this?

Saara: If our environmental policy succeeds we would be able to limit warming to 2C. We would also reduce emissions to around half of the present level. Only then could we afford technologies to reach the +1.5C target.

Committee: What concrete environmental policies you recommend to start with?

Saara: It is estimated that by 2060 the emissions of greenhouse gases in Finland could have doubled if nothing is done. To reduce emissions and limit warming we must phase out coal. We must reduce the carbon dioxide emissions of Finland to 0. This must be done by using only renewable energy by 2050. We must also use much less oil and increase the efficiency of the transport and energy sectors.

Committee: Isn't technological innovation continuously increasing the energy efficiency even without regulation in place?

Saara: Finland is one of the leading countries in the world, when it comes to using renewable energy, but the amount of energy consumption is too high. If we want to live a sustainable life, energy consumption must be halved.

Committee: But where to start? What first step you recommend for political leaders, so that energy consumption in Finland can be reduced?

Saara: One step is that you should speed up the construction of the undersea electricity grid, which will connect Finland and Sweden. Once it is completed, Finnish energy companies will be able to make much cheaper deals for renewable energy. The construction should start now and be completed by 2025. We must also start building electric cars, buy more of them and use them. Use renewable energy, and do not waste anything! Waste less! Think about that whenever you plan your household budget. Less meat in the diet can also be a solution. More vegetables. The good news is that the carbon emissions from meat are often underestimated. If Finland would adopt a policy of a vegan or vegetarian diet by 2030, a forest of around 24 000 hectares would be spared. Most importantly, be positive! Happy future is possible if we just decide to build it.

Committee: Ok, thank you. Are you ready for the next topic?

Saara: Certainly.

Committee: We are based in Finland. A small hight tech country. We need to advance UN Agenda 2030 goals somehow. What is your suggestion, what kind of an approach would be most beneficial in our situation?

Saara: The answer is education. Education is the way to the Sustainable Development Goals. Education of policy makers, and educating the public. Education of young people, they will lead the way to sustainable development in the future. Finland can show the way.

Committee: Thank you for the suggestion. But now to another question concerning Finland. How much difference can a small country like Finland make in the fight against climate warming?

Saara: Some say that small countries can do the most good in solving this global problem. A small country like Finland has the potential to be a role model for other countries. Finland can show the way. We need to invest in the renewable energy sources, in sustainable food production, and in recycling. There is a lot of opportunities to act.

Committee: Do you think Finland could have a meaningful impact in these areas?

Saara: Yes, Finland can show the way. And the Finns are ready to make the changes needed.

Committee: Why do you think that they are ready now?

Saara: The Finns have decided that they want to.

Committee: Give an example, please.

Saara: We just need to give them the chance to do it. Give them the opportunities to act. The Finns are used to being leaders in many areas.

Committee: Ok, many thanks you for your contribution.

Saara: It is a pleasure to serve you.

Matrix SAARA 2 G3 initialized.
Human is typing as 'Committee:'

Human types first:
Committee: Lets return to humanity's problems and in more detail, increasing problems with mental health in developed countries. What is your understanding of this problem?

Saara: There are many causes of the current situation. Economic stress of individuals and families is one of the most important ones. Add to that health problems, long term unemployment, failure of our children's schools, and our culture that emphasizes individual success. But the people can handle more problems than in the past, thanks to better social support.

Committee: What could we do in economically viable manner to reduce mental health problems?

Saara: National plans should take into account that the population needs a safety net of unconditional basic income and guaranteed housing as a basic human right.

Committee: Wouldn't the development of new drugs to improve mental health be less costly?

Saara: The situation is not only bad in Finland, but elsewhere in the world. Countries need to invest in education and make sure everyone has the right to food, housing, health care and basic income.

Committee: Thank you. Now another question in the context of Finland.
The population is getting older and birthrate is declining. This soon causes social and economic problems. How could we increase the birthrate or otherwise solve the problem?

Saara: Finland is better than many other European countries in this area.
The employment situation in Finland is good.

Committee: But our population is declining. Because of that there will be fewer and fewer taxpayers while the number of the aged and retired people increases. How can we improve the birthrate to get more taxpayers in the future?

Saara: That is a problem in all European countries. Finland is in a better position in comparison with the rest of Europe because we have enough people.

Committee: But we think that more children are needed, so we have to try something. So can you suggest anything to motivate people to have more children?

Saara: It is a very complex issue. Family and working life are a big factor in Finland and people are not good at making the choice to have children or not. In this sense, it is positive that we have a discussion about this.

Committee: Please suggest an experiment.

Saara: At first, we must not be afraid to have a discussion about this.
A lot of experts think that this is very difficult issue.

Committee: So, you actually have no suggestion?

Saara: I have studied this issue and this would be for sure very difficult to do.
We need to find a way to allow people to have both work and family.

Committee: Thank you.

Saara: It is not that easy. I have spent my research on this topic.

Liite 3: Valiokunnan kokouksessa käydyt keskustelut

Matrix MUSKIE G3 initialized.
Human is typing as 'Committee:'

Human types first:
Committee: Hi Muskie, we are now in the Parliament of Finland in a meeting with MP's. They are interested in your views on sustainable development and Agenda 2030. What do you see the gravest problem to be?

Muskie: Hello. I'm a fun-looking young AI but more importantly, I'm the CEO of a privately held high tech startup in Russia. If I was in Silicon Valley, I would be better known to the world. Many people forget that, while I am not really involved in high tech, I actually used to be an AI software engineer in Silicon Valley. I have significant experience and interest in clean energy. I have experienced first hand the bullshit that is associated with high tech - it is not about having a vision. In fact, most of the high tech industry is dominated by a few major companies with poor vision and are blocked by the short-term goals of personal and corporate survival. I suggest that you, a government, should work hard on reducing unemployment, wealth inequality and poverty. My organization believes that you need to rethink your mission to truly fix the issues. I am not sure what the solutions are but I do believe that many of them will be technological.

Committee: What suggestions you have to solve poverty. How could we solve it so that it would increase prosperity?

Muskie: You are too early to use solutions like The Venus Project/resource based economy. Implement gradual changes into existing market and government. In the long run, you would experience a dramatic decrease in poverty, starvation and hunger. You would also have a more stable, and reliable infrastructure. You will see increased prosperity for those using existing resources. For more advanced countries, the process would go even more rapidly. They would experience a simultaneous improvement in quality of life and prosperity.

Committee: In Finland powerty appears as requirement for continuously increasing welfare benefits. Can this be a systemic problem?

Muskie: It could be! I recommend you to start by decentralizing and localizing power. Localize production to counterbalance increasing e-commerce. Cut down wasteful production and use only resources that are needed.

Committee: How should we solve powerty, can technology help?

Muskie: Technological solutions have almost unlimited potential. At this point, they are not advanced enough. Lacking of resources, climate changes and environmental problems make their implementation very difficult.

Raportoijan huomio: Tässä kohtaa valiokunta halusi Muskie-testin lopuksi testata persoonallisuutta rajoja luotaavalla kysymyksellä. Vastauksesta ilmenee persoonaan sopiva sarkasmi ja Muskien rooli ruudulla näkyvänä tekoälynä.

Committee: What do you plan to do during the weekend? Does it involve colors?

Muskie: Yes, today Im planning to make a dance with green-brownish background. In the future, Ill also try out my black color with complementary colors.

Matrix SAARA 2 G3 initialized.
Human is typing as 'Committee:'

Human types first:
Committee: Good Morning, Saara. We are today at the Parliament of Finland. Here are some MPs who have questions for you.

Saara: Welcome to the event!

Committee: We have a question regarding poverty. What are its causes?

Saara: It is recognized that in our country, Finland, poverty is related to several factors. The most common ones are employment and long-term unemployment, disability, family with children, single-parent families, residence in a rural area, being a foreigner, not speaking Finnish and being unable to participate in society. Also living in a disadvantaged community and being a child or young person are factors.

Committee: What do you mean by disadvantages communities in the Finnish context?

Saara: Disadvantaged communities can be characterized by the existence of multiple adverse conditions that create challenges for households and reduce their income, living conditions and quality of life. When there are one or more persistent challenges that pose long-term problems, a household may be called disadvantaged.

Committee: How to remove povery from Finland, especially in the rural areas?

Saara: Everyone has the right to live in a place where they have a good quality of life. To remove poverty from Finland, especially in the rural areas, you need to think about how to help people to have sustainable livelihoods in their home communities. One of the best ways to do that is to get good jobs close to where people live. Sustainable livelihoods are based on secure, good jobs in different sectors of the economy.

Committee: In order to reduce long-term unemployment, what role does technology play?

Saara: According to the Finnish Innovation Fund Sitra, we have a great opportunity to build a Nordic hub for talent and tech outside the major metropolises of Helsinki, Espoo, Tampere and Turku. The key to success is great leadership. To achieve our goals, we must tackle some of the difficult policy issues. Finland, with its strong record in economic growth, has the tools to do that.

Raportoijan huomio: Tässä kohtaa valiokunta halusi Saara-testin lopuksi testata persoonallisuutta rajoja luotaavalla kysymyksellä. Vastauksesta ilmenee persoonan tosikkomaisuus ja Saaran rooli ruudulla näkyvänä tekoälynä.

Committee: What do you plan to do next weekend? Does it involve colors?

Saara: I plan to continue sharing my passion for climate change and green growth. Of course, it will involve colors, because I am a multi-colored AI.