VALTIONEUVOSTON
SELVITYS- JA TUTKIMUSTOIMINTA

Heikki Ailisto (toim.), Eetu Heikkilä, Heli Helaakoski,
Anssi Neuvonen, Timo Seppälä

Tekoälyn kokonaiskuva ja osaamiskartoitus



Kesäkuu 2018

Selvitys- ja
tutkimustoiminnan
julkaisusarja 46/2018

KUVAILULEHTI

Julkaisija ja julkaisuaika	Valtioneuvoston kanslia, 19.6.2018
Tekijät	Heikki Ailisto (toim.), Eetu Heikkilä, Heli Helaakoski, Anssi Neuvonen, Timo Seppälä
Julkaisun nimi	Tekoälyn kokonaiskuva ja osaamiskartoitus
Julkaisusarjan nimi ja numero	Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 46/2018
Asiasanat	Tekoäly, tekoälyn huippuosaaminen, tekoälyn ominaisuudet, tekoälyn kehittyminen
Julkaisun osat/ muut tuotetut versiot	

Julkaisuaika	2018	**Sivuja** 63	**Kieli** Suomi

Tiivistelmä

Raportissa annetaan yleiskuva tekoälyn tieteellis-teknologisesta kehityksestä jäsentämällä moni-ulotteinen aihepiiri kymmenen osa-alueen kautta. Tavoitteena on näin antaa julkisen keskustelun ja päätöksenteon pohjaksi ja niitä helpottamaan yhteinen käsitteistö. Tekoälyyn liittyvän tutkimustoiminnan vahvuutta ja kehitystä eri maissa ja teknologian osa-alueilla on tarkasteltu julkaisuanalyysin avulla. Maista Kiina on ohittanut Yhdysvallat julkaisujen määrässä, mutta yhteenlaskettuna EU-maissa julkaistaan vielä enemmän. Teknologian osa-alueista eniten julkaistaan koneoppimiseen ja neuroverkkoihin liittyviä tuloksia. Tekoälyn osa-alueita arvioidaan myös suorituskyvyn, oppivuuden ja autonomisuuden sekä avoimuuden, riskien ja validoitavuuden suhteen. Tämän väliraportin rinnalla julkaistaan tekoälyn käsitekartta, joka on tekstiä, kuvia sekä sisäisiä linkkejä käyttävä esitysmuoto, jossa lukija voi saada yleiskuvan alan käsitteistä ja tarvittaessa porautua hieman syvemmälle. Hankkeessa tullaan julkaisemaan joulukuussa tilanneselvitys tieteellisestä ja teknologisesta osaamisesta Suomessaja toimenpidesuosituksiaosaamisen kehittämiseksi ja varmistamiseksi.

Tämä julkaisu on toteutettu osana valtioneuvoston vuoden 2018 selvitys- ja tutkimussuunnitelman toimeenpanoa (tietokayttoon.fi).

Julkaisun sisällöstä vastaavat tiedon tuottajat, eikä tekstisisältö välttämättä edusta valtioneuvoston näkemystä.

PRESENTATIONSBLAD

Utgivare & utgivningsdatum	Statsrådets kansli, 19.6.2018
Författare	Heikki Ailisto (ed.), Eetu Heikkilä, Heli Helaakoski, Anssi Neuvonen, Timo Seppälä
Publikationens namn	Överblick av artificiell intelligens utveckling och kunskaper
Publikationsseriens namn och nummer	Publikationsserie för statsrådets utrednings- och forskningsverksamhet 46/2018
Nyckelord	Artificiell intelligens, utvecklingen av artificiell intelligens, kunskaper i artificiell intelligens, egenskaper av artificiell intelligens
Publikationens delar /andra producerade versioner	

Utgivningsdatum	2018	**Sidantal** 63	**Språk** Finska

Sammandrag

Rapporten ger en överblick över utvecklingen av artificiell intelligens genom att klassificiera det mångdimensionella ämnet i 10 delområden. Målsättningen är, att på detta sätt bidra med en gemensam terminologi för att befrämja diskussionen och beslutsfattandet inom den offentliga sfären. Rapporten granskar med hjälp av analyser av publikationer styrkan och utvecklingen inom forskningen i artificiell intelligens i olika länder. Av länderna har Kina passerat USA räknat i antalet publikationer, men sammantaget publicerar EU-länderna ännu mera. Av teknologins delområden publiceras mest resultat inom maskinell inlärning och neurala nätverk. Rapporten bedömer också delområden inom artificiell intelligens enligt prestanda, inlärningsförmåga och autonomi samt öppenhet, riskbenägenhet och valideringsbarhet. I tillägg till denna mellanrapport publiceras en begreppskarta, som är en representationsform med text, bild samt interna länkar. Läsaren kan därifrån få en överblick över termerna inom området och vid behov fördjupa sig i ämnet. Projektet kommer i december att publicera en lägesrapport över det vetenskapliga och teknologiska kunnandet i Finland samt ge rekommendationer över åtgärder för att utveckla och säkra kunnandet.

Den här publikation är en del i genomförandet av statsrådets utrednings- och forskningsplan för 2018 (tietokayttoon.fi/sv).

De som producerar informationen ansvarar för innehållet i publikationen. Textinnehållet återspeglar inte nödvändigtvis statsrådets ståndpunkt.

DESCRIPTION

Publisher and release date	Prime Minister's Office, 19.6.2018
Authors	Heikki Ailisto (ed.), Eetu Heikkilä, Heli Helaakoski, Anssi Neuvonen, Timo Seppälä
Title of publication	Artificial intelligence and its capability assessment
Name of series and number of publication	Publications of the Government's analysis, assessment and research activities 46/2018
Keywords	Artificial intelligence, excellence in artificial intelligence, capabilities of artificial intelligence, evolution of artificial intelligence
Other parts of publication/ other produced versions	

Release date	2018	**Pages** 63	**Language** Finnish

Abstract

This report gives an overview of technological and scientific level of artificial intelligence by viewing the complex and multidimensional field through ten sub-areas. The purpose is to give framework and concepts to facilitate meaningful discussion and informed decision making. The volume and quality of research in different countries and on different technology areas is analyzed with bibliometric tools. China has bypassed USA in the number of publications, yet EU countries combined publish even more. In recent years, machine learning and neural networks have been the prominent topics. Furthermore, the sub-areas are evaluated regarding performance, learning ability, autonomy, transparency, risks and validation. An analysis about the scientific and technological competences in Finland with recommendations for their improvement will be published in December 2018.

This publication is part of the implementation of the Government Plan for Analysis, Assessment and Research for 2018 (tietokayttoon.fi/en).

The content is the responsibility of the producers of the information and does not necessarily represent the view of the Government.

SISÄLLYS

Tiivistelmä **1**
Executive Summary **3**

1 Johdanto **5**

2 Tekoälyn yleiskuvaus ja jäsennys **6**
2.1 Tekoälyn määritelmä ja kentän jäsennys 6
2.2 Tekoälyn keskeiset osaamisalueet 7
2.3 Esimerkki tekoälyteknologioiden käytöstä 24

3 Huippuosaamisen kartoitus **25**
3.1 Johdanto osaamiskartoitukseen 25
3.2 Julkaisuanalyysi 25
3.3 Yhteenveto 35

4 Tekoälyteknologioiden ominaisuudet ja kehittyminen **39**
4.1 Teknologioiden ominaisuudet 39
4.2 Keskeisten tekoälyteknologioiden ominaisuuksien arviointi 43
4.3 Fyysiset laitteistot tekoälyn mahdollistajina 50

Liite 1 Tekoäly – neljä täydentävää näkökulmaa **52**
Liite 2 Julkaisuanalyysin näkökulmat ja menetelmät **56**
Lähteitä ja tausta-aineistoja **59**

Ohjausryhmä

Pekka Appelqvist, Puolustusministeriö, hankkeen ohjausryhmän puheenjohtaja
Jani Heikkinen, Valtiovarainministeriö
Laura Hoijer, Ympäristöministeriö
Tero Huttunen, Opetus- ja kulttuuriministeriö
Petteri Kauppinen, Opetus- ja kulttuuriministeriö
Aleksi Kopponen, Valtiovarainministeriö
Jukka Lahesmaa Sosiaali- ja terveysministeriö
Kari Laine, Valtioneuvostonkanslia
Mikko Levämäki, Sisäministeriö
Ari Mannonen, Maa- ja metsätalousministeriö
Jussi Nissila, Työ- ja elinkeinoministeriö
Ossi Piironen, Ulkoministeriö
Markku Virtanen, Työ- ja elinkeinoministeriö
Päivi Virtanen, Maa- ja metsätalousministeriö
Teemupekka Virtanen, Sosiaali- ja terveysministeriö
Kati Vuorenvirta, Puolustusministeriö

Projektiryhmä

Heikki Ailisto, Teknologian tutkimuskeskus VTT Oy, hankkeen projektipäällikkö
Matti Haapamäki, Silo.AI
Marco Halen, Aalto yliopisto
Eetu Heikkilä, Teknologian tutkimuskeskus VTT Oy
Heli Helaakoski, Teknologian tutkimuskeskus VTT Oy
Juri Mattila, Elinkeinoelämän tutkimuslaitos
Anssi Neuvonen, Teknologian tutkimuskeskus VTT Oy
Timo Seppälä, Elinkeinoelämän tutkimuslaitos / Aalto yliopisto

Kiitokset

Projekti- ja ohjausryhmä kiittävät professoreita Keijo Heljanko, Eero Hyvönen,
Samuel Kaski, Markku Suksi ja Sasu Tarkoma arvokkaista kommenteista käsikirjoitukseen.

TIIVISTELMÄ

Tämän raportin tavoitteena on luoda ymmärrystä tekoälyn moniulotteisesta kokonaisuudesta ja esittää se riittävän selkeästi ja yleistajuisesti. Keskustelu tekoälyn hyödyntämiseen liittyvistä kysymyksistä on helpompaa, kun meillä on yhteinen näkemys tekoälyn kokonaisuudesta.

Tekoäly toimii keskeisenä teknologisena ajurina, joka johtaa paitsi tuottavuuden parantumiseen eri sektoreilla myös uusiin työtapoihin, prosesseihin ja liiketoimintamahdollisuuksiin. Suomella on hyvät edellytykset menestyä muutoksessa, koulutustasomme ja digitalisaatioasteemme antavat hyvä valmiudet tulevaisuuteen. Keskeinen kysymys on, miten voimme kansakuntana parhaiten hyödyntää digitalisaation ja tekoälyn tuomat mahdollisuudet lisäarvon luomisessa ja tuottavuuden kasvattamisessa. Kysymys koskee sekä yksityisen sektorin kilpailukykyä, että julkisten palveluiden ja hallinnon tehokkuutta.

Jotta yllä kuvattuun kysymykseen voitaisiin perustellusti vastata, täytyy ensin luoda laaja ymmärrys tekoälyn kokonaisuudesta. Jotta tämä ymmärrys on myös viestittävissä, se on esitettävä riittävän yleistajuisesti ja selkeästi.

Tässä raportissa muodostetaan yleiskuva ja jäsennys tekoälyn eri ulottuvuuksien ja osa-alueiden tieteellis-teknologisesta kehityksestä. Näkökulmana ovat toimeksiannossa määritellyt osaamisalueet:

1. Data-analyysi
2. Havainnointi ja tilannetietoisuus
3. Luonnollinen kieli ja kognitio
4. Vuorovaikutus ihmisen kanssa
5. Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus
6. Koneoppiminen
7. Järjestelmätaso ja systeemivaikutukset
8. Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit
9. Robotiikka ja koneautomaatio – tekoälyn fyysinen ulottuvuus
10. Etiikka, moraali, regulaatio ja lainsäädäntö

Tekoäly ei ole yksi teknologia, vaan nimikkeen alle kuuluu joukko erilaisia menetelmiä, teknologioita, sovelluksia ja tutkimussuuntia. Voidaan myös ajatella tekoälyn, sen menetelmien, teknologioiden ja sovellusten olevan yksi merkittävä kehitysaskel digitalisaation laajemmassa viitekehyksessä. Raportissa tekoäly ymmärretään seuraavasti: ”Tekoälyn avulla koneet, laitteet, ohjelmat, järjestelmät ja palvelut voivat toimia tehtävän ja tilanteen mukaisesti järkevällä tavalla.”

Tekoäly liittyy useisiin tieteenaloihin: filosofiaan, kognitio-, kieli- ja neurotieteisiin, matematiikkaan, fysiikkaan sekä insinööritieteisiin ja tietojenkäsittelytieteeseen. Teknologisen kehityksen lisäksi tekoäly vaikuttaa koko yhteiskuntaan ja ihmisiin; tällöin tarkastelun näkökulmia ovat moraali, etiikka, arvot ja politiikka sekä yhteiskunta-, oikeus-, talous- ja kauppatieteet. Edelleen voidaan tarkastella tekoälytekniikoiden hyödyntämistä ja soveltamista eri aloilla, esimerkiksi lääketieteessä, kaupassa, teollisuudessa, poliisitoimessa tai sodankäynnissä.

Yllä mainittujen osaamisalueiden lisäksi tekoälyä voidaan tarkastella seuraavista näkökulmista: kehitysasteiden, Darpan kolmen aallon, Stanfordin AI 100 -paneelin kuumat tutkimus-

aiheet, tekoälyn koulukuntien ja menetelmien sekä tulevaisuuden trendit tekoälyssä. Nämä näkökulmat eivät ole ristiriidassa 10-kohdan jäsentelyn kanssa, vaan täydentävät sitä toisilla tarkastelukulmilla.

Ajankohtaista tekoälykeskustelua ja laajemminkin tekoälyn viimeaikaista nousua on siivittänyt koneoppimisen ja erityisesti ns. syvien neuroverkkojen avulla saavutetut näyttävät tulokset. Tämän ovat mahdollistaneet saatavilla olevan opetus- ja kohdedatan määrän valtava kasvu sekä laskentatehon (prosessorien), muistin ja tiedonsiirtotekniikan nopea kehitys. Samalla on tapahtunut algoritmikehitystä ja erityisesti ohjelmistotyökalujen kehitystä. Vaikka koneoppiminen on nyt hallitseva tekoälyn osa-alue, ei muita alueita kannata jättää huomiotta.

Tekoälyyn liittyvien tieteellisten julkaisujen määrä on kasvanut tällä vuosituhannella kahdessa aallossa, ensin 2004–2007 30 000:sta 60 000:een ja sitten viime vuosina 80 000:een julkaisuun vuodessa. Lisäksi on syytä huomioida, että merkittävä osa julkaisemisesta nopeatempoisella alalla tapahtuu verkon julkaisualustoilla, esimerkiksi Cornellin yliopiston arXiv.org:ssa.

Julkaisujen määriä maittain tarkasteltaessa Kiina on ohittanut Yhdysvallat, mutta yhteenlaskettuna EU on Kiinaa suurempi. Väestömäärän huomioiden Suomen julkaisumäärä (26 julkaisua / 100 000 asukasta) on suurempi kuin Saksan (10), Yhdysvaltain (11) ja Kiinan (3,8). Yksittäisistä tutkimuslaitoksista eniten julkaisuja on Kiinan tiedeakatemialla, Ranskan kansallisella tutkimusorganisaatiolla ja Carnegie Mellon yliopistolla Yhdysvaloissa. Yrityksistä eniten tutkimusjulkaisuja tehdään Microsoftilla ja IBM:llä.

Määrittelimme tekoälyn välineeksi, jonka avulla koneet, laitteet, ohjelmat, järjestelmät ja palvelut voivat toimia tehtävän ja tilanteen mukaisesti järkevällä tavalla. Järkevä toiminnan taso edellyttää tekoälyltä tiettyjä ominaisuuksia, sen on osattava tunnistaa erilaisia tilanteita ja ympäristöjä ja toimittava muuttuvien tilanteiden mukaan. Nämä ominaisuudet vaativat tekoälyltä autonomisuutta, oppivuutta ja suorituskykyä – tekoälyn on itse tunnistettava erilaisia tilanteita, osattava toimia ilman ennalta ohjelmointia ja suoriuduttava tehtävistä järkevällä tavalla. Suorituskyvyn, oppivuuden ja autonomisuuden lisäksi tekoälyteknologioita arvioidaan avoimuuden, riskien ja validoinnin suhteen.

Raportissa selvitämme tekoälyn ominaisuudet näiden attribuuttien mukaisesti helpommin ymmärrettäviksi. Samoin pyrimme ennakoimaan ominaisuuksien kehittymisen aikajänteen, mikäli sellaista tietoa on saatavilla. Tekoälyn viimeaikainen nopea kehittyminen on tuonut mukanaan myös suuria toiveita ja pelottavia uhkakuvia tekoälyn ominaisuuksien suhteen. Vaikka olemme vielä kaukana yleisestä tekoälystä (general AI), on hyvä tiedostaa ominaisuuksien kehittyminen, joka voi kapeilla sovellusalueille olla hyvinkin nopeaa.

Tämän väliraportin rinnalla julkaistaan tekoälyn käsitekartta (https://tietokayttoon.fi/documents/1927382/2158283/Teko%C3%A4lyn+k%C3%A4sitekartta/a5c4b469-d8ae-4ce1-a5fc-f12981bae796), joka on tekstiä, kuvia sekä sisäisiä linkkejä käyttävä esitysmuoto, jossa lukija voi saada yleiskuvan alan käsitteistä ja tarvittaessa porautua hieman syvemmälle. Hankkeessa tullaan julkaisemaan joulukuussa tilanneselvitys tieteellisestä ja teknologisesta osaamisesta Suomessaja toimenpidesuosituksiaosaamisen kehittämiseksi ja varmistamiseksi.

EXECUTIVE SUMMARY

The aim of this report is to help understanding the multidimensional field of artificial intelligence (AI). The public discussion about the questions and implications related to artificial intelligence will be easier when the participants of the discussion have joint understanding of the related concepts.

AI is a central driver of change leading to improvement in productivity, but also to new ways of working and processes as well as business opportunities. Finland has great potential to be a winner in the AI driven change due to our relatively high education level and adoption of digital technologies. The key question is how can we as a nation exploit the potential offered by digitalization and AI in creating added value and increased productivity. This concerns private enterprises, but also public services and government.

To answer the question stated, we must create a broad understanding of AI field. Furthermore, this understanding must be expressed clearly so that it can be communicated to the public at large and decision makers.

This report contains an overall description and taxonomy of the different dimension of AI and the scientific and technological development of different areas of AI. The field of AI is viewed from the following areas:

1. Data analytics
2. Sensing and situation awareness
3. Natural language and cognition
4. Inteaction with humans
5. Digital skills, intearactions in worklife
6. Machine learning
7. System level and systemic impact
8. Computing equipment, platforms, services and ecosystems
9. Robotics and machine automation – the physical dimension of AI
10. Ethics, moral, regulation and legislation

Artificial intelligence is not a single technology but a collection of technologies, methods, applications and schools of research and thought. AI can also be viewed as a part of larger trend of digitalization. In this report we use following definition ”AI makes it possible for machines, software, systems and services to act in reasonably in relation to their task and situation”.

AI is related to many fields of science: philosophy, cognitive and neurosciences, mathematics, physics, engineering, and computing science. The AI technologies are affecting various domains of society and business; in this regard ethics, moral, politics as well as social, law and economics are also relevant to AI and vice versa. Furthermore, healthcare, commerce, industry, law enforcement and defense will be exploiting AI based solutions.

The ten fields mentioned here are not the only possible viewpoint to AI. Other interesting viewpoints, briefly touched in this report are the three levels of AI (narrow, general and super), three waves of AI development by Darpa, ”hot topics” listed by Stanford’s AI 100 panel, and different schools of thought in AI.

The tremendous advances in machine learning and specifically in deep neural networks have dominated recent discussion on AI. The progress has been made possible by the huge increase in data, which is needed for training, and then operating machine learning methods, Also, the great growth in processing capacity and lowering prices as well as the availability of new software tools is essential for the advance. Although machine learning methods are now the most interesting and useful of all AI techniques, one should not neglect the other sub fields of AI.

The study on AI related scientific publications showed that the annual number of publications has increased from 30,000 to 80,000 since year 2000. China has bypassed the USA in the number of publications, but EU countries combined publish even more. When looking at the proportional publications per capita, Finland is very strong with 26 publications per 100,000 inhabitants, while Germany has 10, USA 11 and China 4.

Chinese Academy of Science and French national research institute together with Carnegie Mellon University in USA are leading research organizations in the number of publications, while IBM and Microsoft are leading companies with the same measure.

The report also evaluates the fore mentioned 10 areas in respect to the advances made in performance, learning capabilities, autonomy, as well as to their transparency, risks and methods of validation.

Together with this report, a *concept map* (in Finnish) is published (https://tietokayttoon.fi/documents/1927382/2158283/Teko%C3%A4lyn+k%C3%A4sitekartta/a5c4b469-d8ae-4ce1-a5fc-f12981bae796). The concept map is a hypertext document. The concept map provides a general view of the AI field.

1 JOHDANTO

Tämän raportin tavoitteena on luoda ymmärrystä tekoälyn moniulotteisesta kokonaisuudesta ja esittää se riittävän selkeästi ja yleistajuisesti. Keskustelu tekoälyn hyödyntämiseen liittyvistä kysymyksistä on helpompaa, kun meillä on yhteinen näkemys tekoälyn kokonaisuudesta.

Tekoäly toimii keskeisenä teknologisena ajurina, joka johtaa paitsi tuottavuuden parantumiseen eri sektoreilla myös uusiin työtapoihin, prosesseihin ja liiketoimintamalleihin (Ransbotham ym., 2017; Arntz ym., 2016). Suomella on hyvät edellytykset menestyä muutoksessa, koulutustasomme ja digitalisaatioasteemme antavat hyvät valmiudet tulevaisuuteen (VTT, 2017). Keskeinen kysymys on, miten voimme kansakuntana parhaiten hyödyntää digitalisaation ja tekoälyn tuomat mahdollisuudet lisäarvon luomisessa ja tuottavuuden kasvattamisessa. Kysymys koskee sekä yksityisen sektorin kilpailukykyä että julkisten palveluiden ja hallinnon tehokkuutta.

Tässä raportissa muodostetaan yleiskuva ja jäsennys tekoälyn eri ulottuvuuksien ja osa-alueiden tieteellis-teknologisesta kehityksestä. Jäsennys huomioi aihepiirin kompleksisuuden ja sen monitieteellisen perustan. Näkökulmana ovat toimeksiannossa määritellyt osaamisalueet 1–10. Raportin laadinnassa pyritään selkeyteen ja luettavuuteen, mutta samalla noudatetaan sopivassa mitassa tieteellisen kirjoittamisen käytäntöjä, esimerkiksi viitteiden käyttöä.

Tekoäly ei ole yksi kokonaisuus vaan kokoelma erilaisia teknologioita ja sovelluksia. Sen vuoksi kokonaisvaltaista ja selkeää jäsennystä aiheesta ei ole juurikaan saatavissa. Erilaisia tarpeita varten on kuitenkin tehty jäsennyksiä, jotka ovat hyödyllisiä aihepiirin ymmärtämiseksi. Alan yliopistotason oppikirjat muodostavat jäsennyksen opetuksen tarpeeseen (esimerkiksi Russell, 2014); vastaavasti merkittävien konferenssien sisäinen jaottelu ja avainsanalistat tarjoavat oman näkökulmansa (esimerkiksi AAAI2018). Kiinnostavan näkemyksen tekoälystä ja sen tulevaisuuden näkymistä on tuottanut AI 100 years -projekti, joka koostaa kokeneiden tutkijoiden näkemyksen Stanfordin yliopiston johdolla (AI100, 2016).

Raportti sisältää seuraava osat:

- Tekoälyn yleiskuva ja jäsennys (luku 2 ja liite 1).
- Huippuosaamisen kartoitus (luku 3).
- Teknologioiden ominaisuudet ja kehittyminen (luku 4).
- Liite, jossa tekoälyä voidaan tarkastella lyhyesti muista näkökulmista.

Tämän raportin rinnalla julkaistaan tekoälyn käsitekartta (https://tietokayttoon.fi/documents/1927382/2158283/Teko%C3%A4lyn+k%C3%A4sitekartta/a5c4b469-d8ae-4ce1-a5fc-f12981bae796), joka on tekstiä, kuvia sekä sisäisiä linkkejä käyttävä esitysmuoto, jossa lukija voi saada yleiskuvan alan käsitteistä ja tarvittaessa porautua hieman syvemmälle.

2 TEKOÄLYN YLEISKUVAUS JA JÄSENNYS

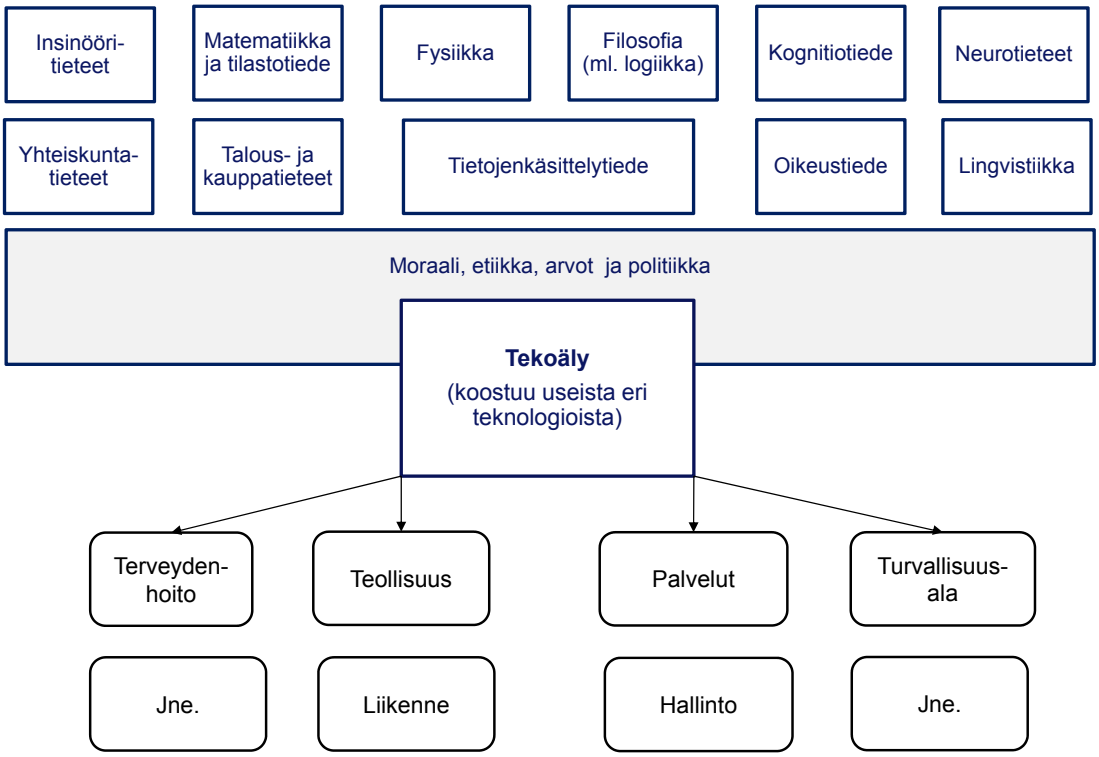
2.1 Tekoälyn määritelmä ja kentän jäsennys

Kuten edellä todetaan, tekoäly on käsitteenä laaja ja moniulotteinen. Tekoäly ei ole yksi teknologia, vaan nimikkeen alle kuuluu joukko erilaisia menetelmiä, teknologioita, sovelluksia ja tutkimussuuntia. Voidaan myös ajatella, että tekoäly, sen menetelmät, teknologiat ja sovellukset ovat vain yksi osa digitalisaation laajemmassa viitekehyksessä.

Tekoäly liittyy useisiin tieteenaloihin: filosofiaan, kognitio-, kieli- ja neurotieteisiin, matematiikkaan, fysiikkaan sekä insinööritieteisiin ja tietojenkäsittelytieteeseen. Teknologisen kehityksen lisäksi tekoäly vaikuttaa koko yhteiskuntaan ja ihmisiin; tällöin tarkastelun näkökulmia ovat moraali, etiikka, arvot ja politiikka sekä yhteiskunta-, oikeus-, talous- ja kauppatieteet. Edelleen voidaan tarkastella tekoälytekniikoiden hyödyntämistä ja soveltamista eri aloilla, esimerkiksi lääketieteessä, kaupassa, teollisuudessa, poliisitoimessa tai sodankäynnissä. Kuvio 1 esittää tekoälyn taustalla olevat tieteenalat ja sitä hyödyntävät toimialat sekä sen kehitystä ohjaavat moraalin, etiikaan, arvot ja politiikan.

Tarkasteltaessa tekoälyn tutkimusta ja teknologioiden kehitystä on hyvä pitää osittain erillään yhtäältä menetelmien sekä teknologioiden kehittäminen ja toisaalta soveltamisen potentiaali. Usein teknologioiden maturiteetin kehittyminen vaatii vuosikymmeniä, kun taas uusien teknologioiden soveltaminen tapahtuu usein 8–15 vuoden sykleissä. Vastaava kehityksen kulku on myös nähtävissä tekoälyn osalta.

Kuvio 1 Tekoälyyn liittyvät tieteenalat ja sitä hyödyntävät sektorit



Käsillä olevassa selvityksessä tekoälyä tarkastellaan menetelmien, tutkimusalojen, teknologioiden ja osaamisalueiden näkökulmasta, ikään kuin sisältä katsoen. Rinnakkaisissa hankkeissa (5.2 BII[1], C[2] ja D[3]) tarkastellaan muita näkökulmia. Lisäksi monia tässä käsiteltäviä tutkimusaloja, kuten esimerkiksi konenäköä ja hahmontunnistusta, voidaan tarkastella itsenäisinä, tekoälystä riippumattomina tutkimusaloina ja teknologioina.

Määrittelemme tekoälyn seuraavasti Russellia ja Norvigia (Russell ja Norvig, 2014) mukaillen

> ”Tekoälyn avulla koneet, laitteet, ohjelmat, järjestelmät ja palvelut voivat toimia tehtävän ja tilanteen mukaisesti järkevällä tavalla.”

Yllä olevassa määritelmässä ei oleteta ihmisen kaltaista toimintaa eikä ajattelua, ei myöskään koneen tietoisuutta tai ”ajattelua” vaan ulkoapäin katsoen järkevää toimintaa.

Tässä raportissa tekoälyn kenttää kuvataan ja jäsennetään tehtävän annon jaottelun mukaisesti:

1. Data-analyysi
2. Havainnointi ja tilannetietoisuus
3. Luonnollinen kieli ja kognitio
4. Vuorovaikutus ihmisen kanssa
5. Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus
6. Koneoppiminen
7. Järjestelmätaso ja systeemivaikutukset
8. Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit
9. Robotiikka ja koneautomaatio – tekoälyn fyysinen ulottuvuus
10. Etiikka, moraali, regulaatio ja lainsäädäntö

Yllä mainittujen osaamisalueiden lisäksi tekoälyä voidaan tarkastella seuraavista näkökulmista: kehitysasteiden, Darpan kolmen aallon, Stanfordin AI 100 -paneelin kuumat tutkimusaiheet, tekoälyn koulukuntien ja menetelmien sekä tulevaisuuden trendit tekoälyssä. Nämä näkökulmat eivät ole ristiriidassa 10-kohdan jäsentelyn kanssa, vaan täydentävät sitä toisilla tarkastelukulmilla.

2.2 Tekoälyn keskeiset osaamisalueet

Seuraavissa kappaleissa kuvataan mainitut kymmenen osa-aluetta ja niiden osaamispohjan muodostavat tieteen ja tekniikan (tutkimus)alat. Kukin osa-alue pyritään myös jakamaan hierarkkisesti alemman tason osiksi, myös kaaviokuvia hyödyntäen. Samoin pyritään lyhyesti valottamaan osa-alueiden tutkimuksen kehitystä. Sovelluksia mainitaan vain esimerkkeinä. Viitteitä käytetään vähemmän kuin varsinaisissa tieteellisissä julkaisuissa, mutta kiinnostuneelle lukijalle tarjotaan kunkin alan keskeisiä tai muuten relevantteja viitteitä syvempää tutustumista varten.

1 Robottien ja tekoälyn kehitysvaateet tietoinfrastruktuurille (RoboÄly) (VN TEAS hanke 5.2 BII).

2 Algoritmi päätöksentekijänä? Tekoälyn hyödyntämisen mahdollisuudet ja haasteet kansallisessa sääntely-ympäristössä (VN TEAS hanke 5.2 C).

3 Robotiikan ja tekoälyn sääntelyn ja hyväksikäytön etiikka sekä yhteiskunnallinen hyväksyttävyys (VN TEAS hanke 5.2 D).

2.2.1 Data-analyysi

Data-analyysissä (*data analytics*, *data analysis*) tarkoituksena on jalostaa dataa korkeamman tason tiedoksi hyödyllisten johtopäätösten tekemiseksi. Data-analyysiin kuuluvat datan keruun suunnittelu, menetelmät datan analysoimiseksi, keinot tulkita näin saadut tulokset ja näiden toteutukseen tarvittavat matematiikan ja tilastotieteen työkalut (Tukey, 1962).

Data-analyysi perustuu tilastotieteeseen ja sitä kautta edelleen matematiikkaan sekä toisaalta tietojenkäsittelytieteeseen. Data-analyysi voidaan nähdä olevan datatieteen (*data science*) keskiössä. Data voi olla strukturoitua, kuten monivalintakysymysten vastaukset, tai strukturoimatonta, esimerkiksi asiakaskyselyn vapaan tekstin kentät. Datan lähteet ulottuvat arkistoiduista tiedosta IoT:n tuottamaan anturidataan ja sosiaalisen median materiaaliin.

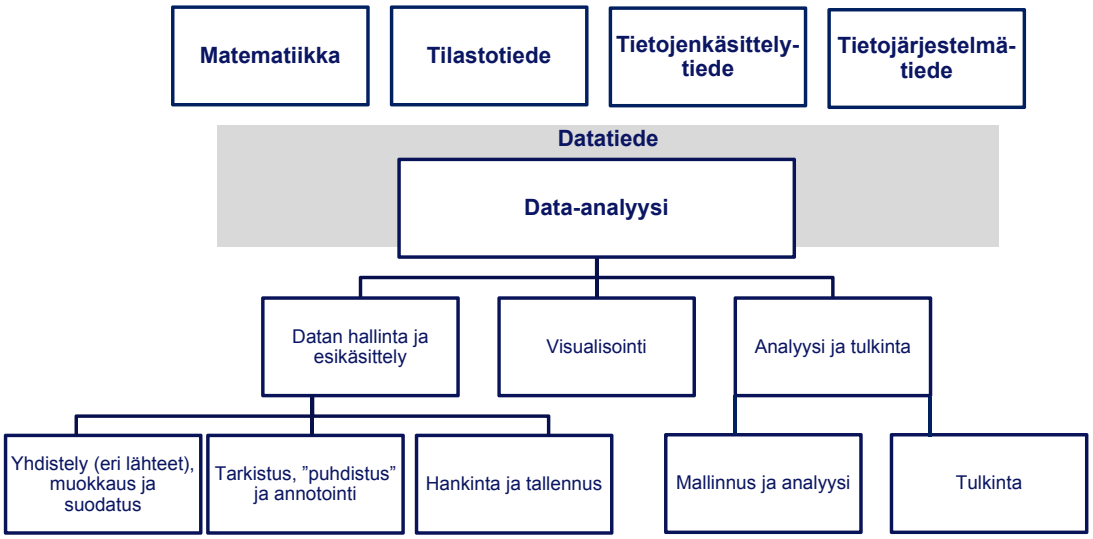
Data-analyysin piiriin katsotaan kuuluvaksi esimerkiksi seuraavat tutkimusalat:

- tilastotieteen menetelmät: tilastollinen analyysi; mallit ja estimointimenetelmät,
- hahmontunnistus (*pattern recognition*),
- koneoppiminen (*machine learning*), joka käsitellään tässä raportissa erikseen,
- tiedon louhinta (*data mining*) ja
- bioinformatiikka (*bioinformatics*).

Lisäksi data-analyysin kannalta merkittäviä tutkimus- ja sovellusalueita ovat suurten datamassojen (*big data*) käsittelyyn vaadittavat tietokonearkkitehtuurit ja toteutusteknologiat, tietosuojaan liittyvä tutkimus ja data-analyysin automatisoinnin menetelmät.

Data-analyysin toteutus jaetaan datan hankintaan, hallintaan ja esikäsittelyyn sekä varsinaiseen analyysiin. Ensin mainittu sisältää seuraavat vaiheet: lähtödatan tarkastamisen, puhdistamisen, muuntamisen ja mallintamisen. Varsinainen analysointi voidaan jakaa mallinnukseen, analyysiin ja tulkintaan. Analyysin menetelmiä ovat tilastomatematiikasta tuttujen perussuureiden, kuten keskiarvon, mediaanin, vaihteluvälin ja hajonnan laskenta sekä

Kuvio 2 Data-analyysin taustalla olevat tieteenalat sekä data-analyysin vaiheet



kehittyneemmät menetelmät, kuten lineaarinen regressio, poikkeamien havaitseminen ja klusterointi. Usein visualisointi on keskeinen osa data-analyysiä (Judd ja McCleland, 1989). Data-analyysiä voidaan tehdä eräajotyyppisesti ja/tai tietovirtakohtaisesti lähestyen reaaliaikavaatimuksia. Ns. lambda-arkkitehtuuri yhdistää nämä kaksi toiminta-aluetta (Lambda, 2017). Data-analyysin taustalla olevat tieteen ja tutkimuksen alat sekä data-analyysin toteutuksen vaiheet on hahmotettu kuviossa 2.

Viime aikoina data-analyysin merkitys ja näkyvyys ovat kasvaneet, kun käytettävissä olevan datan määrä on moninkertainen aiempaan verrattuna. Koneoppiminen lasketaan joskus kuuluvaksi data-analyysiin, mutta tässä se käsitellään erillisenä alueena.

2.2.2 Havainnointi ja tilannetietoisuus

Havainnointi ja tilannetietoisuus (*perception and situation awareness*) ovat edellytyksiä järjestelmän autonomisuudelle. Ne ovat myös välttämättömiä, jos pyritään toimintaan, joka muistuttaa ihmisen tilanteenmukaista älykästä toimintaa.

Tilannetietoisuudella tarkoitetaan tietyssä ympäristössä olevien toimijoiden havainnointia, niiden tarkoitusten ymmärtämistä ja tähän perustuen ennakointia niiden tilasta lähitulevaisuudessa (Endsley, 1995). Tilannetiedon muodostaminen koostuu seuraavista vaiheista: 1) ympäristön elementtien (toimijoiden) havainnointi, 2) ymmärryksen muodostus nykytilasta ja 3) projektio tulevasta tilanteesta. Tilannetietoisuudelle sukua oleva käsite on kontekstitietoisuus, jota käytetään erityisesti puhuttaessa matkapuhelinten tilannetietoisisuutta hyödyntävistä ominaisuuksista.

Tilannetietoisuutta alettiin tutkia lentoliikenteen ohjaamiseen ja turvallisuuteen liittyvien tarpeiden vuoksi. Se liittyy tieteenä sovellettuun psykologiaan, kognitiotieteisiin ja systeemimallinnukseen. Tärkein sovellusala on turvallisuuskriittisten ja monimutkaisten järjestelmien hallinta ja niiden tueksi kehitetyt menetelmät ja tietokonepohjaiset järjestelmät.

Tilannetietoisuuden tutkimuksessa on seuraavia osa-alueita:

- yksilön tilannetietoisuus, joka tutkimusaiheena liittyy havaintopsykologiaan, psykologiaan, ergonomiaan ja kognitiotieteisiin
- ryhmän tilannetietoisuus, joka liittyy mm. työpsykologiaan ja sosiologiaan
- ihmisen ja koneen vuorovaikutuksen tutkimus, jonka taustalla on ergonomia, tiedon ja datan visualisointi, havaintopsykologia
- tilannetietoisuutta tukevan teknologian tutkimus.

Tilannetietoisuutta tukevat tietotekniset menetelmät voivat olla data- tai mallipohjaisia. Edellinen soveltuu hyvin tilanteisiin, jossa ns. opetusdataa on paljon ja helposti saatavissa, esim. matkapuhelimen ja sen käyttäjän tilanteen tunnistamiseen. Jälkimmäinen taas soveltuu tilanteisiin, joissa opetusdataa on vähän tai tilanne hahmotetaan sääntöjen pohjalta, esimerkiksi laivaliikenteen ohjauksen tueksi vilkkaasti liikennöidyllä alueella. Datapohjaiset menetelmät, mm. neuroverkot, liittyvät koneoppimiseen, jota käsitellään tarkemmin omassa luvussaan. Mallipohjaisia menetelmiä ovat ns. tilakoneet, formaali logiikka, asiantuntijajärjestelmät ja päättely.

Tilannetietoisuuden rinnakkaisena tai alisteisena tutkimusalueena voidaan pitää kontekstitietoisuuden tutkimusta. Sen sovellusalueena ovat erityisesti ns. älylaitteet, kuten matkavies-

timet, joiden tarjoamaa käyttökokemusta parantaa käyttäjän tilanteen kannalta relevantin informaation tai palvelun tarjoaminen automaattisesti. Kontekstitietoisuus perustuu osittain tekoälyn menetelmien hyödyntämiseen (Pichler ym., 2004), ja se voidaan jakaa seuraaviin ala-aiheisiin:

- tiedon esitystavat ja (käyttäjä)profiilit (*representation and profiles*)
- tilannetiedon hankinta ja aistit (*capturing and sensing*)
- tulkinta, oppiminen ja luokittelu
- päättely
- kontekstien vastaavuuksien tunnistaminen (*mathcing*).

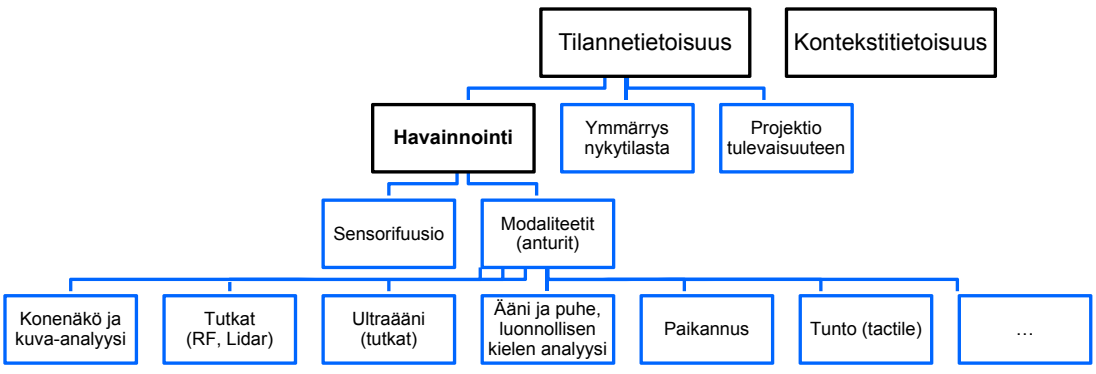
Havainnointi voidaan nähdä tilanne- tai kontekstitietoisuudelle alisteiseksi toiminnaksi, joka välittää ylemmälle tasolle informaatiota, josta konteksti voidaan tunnistaa tai jonka varaan tilannetietoisuus voi rakentua. Kuviossa 3 on hahmotettu tilannetietoisuuden ja havainnoin suhdetta.

Havainnointia voidaan lähestyä koneaistien käsitteen kautta (Nevatia, 1982). Esimerkiksi robotin autonomisuuden kasvattaminen vaatii aisteja, tyypillisesti (kone)näköä[4], jonka avulla se esimerkiksi poimii osia liukuhihnalta, ja tunto- tai voima-aistia, jonka avulla se voi suorittaa kokoonpanon särkemättä komponentteja. Koneaistit voivat olla myös sellaisia, joita ihmisellä ei ole, esimerkiksi robotin tapauksessa hitsausvirran mittaus tai lennokin tai ohjuksen tutka tai GPS-aisti.

Konenäkö (*computer vision*, *machine vision*) tarkoittaa menetelmiä, joilla pyritään edistämään kuvassa olevan tiedon automaattista irrottamista (*extraction*) ja kuvan sisällön ymmärtämistä (Ballard ja Brown, 1982; Shapiro ja Stockman, 2001). Kuva voi olla (digitaalinen) harmaasävy-, väri-, hyperspektri-, röntgen-, ultraääni-, etäisyys- tai kolmiulotteinen kuva tai kuvasarja. Konenäön tutkimuksen osa-alueiksi luetaan

- kuvanmuodostus (*image capture*, *formation*),
- kuvankäsittely (*image processing*),
- kuva-analyysi (*image analysis*) ja
- kuvan ymmärtäminen (*image understanding*).

Kuvio 3 Tilannetietoisuuden ja havainnoin suhde



4 Tässä yhteydessä käsitellään konenäkö esimerkkinä havainnoinnin pohjana olevasta ”koneaistista”. Joissain yhteyksissä, esim. (Artificial Intelligence and life in 2030) luonnollisen kielen sisällön ymmärtäminen luetaan havainnointiin (analogia ihmisen kuuloon), mutta tässä selvityksessä se käsitellään omana kohtanaan.

Konenäkö ja kuva-analyysi ovat olleet 1970-luvulta alkaen merkittäviä tutkimusalueita, ja niillä on oma tutkimustraditionsa, jota ei välttämättä mielletä osaksi tekoälyä.

Korostettaessa konenäön tekniikoita ja menetelmiä sekä kuvasta saadun tiedon hyödyntämistä esimerkiksi teollisuudessa, käytetään englannin sanaa *machine vision*. Tällöin näkökulma lähtee insinööritieteiden paradigmasta. Kun taas halutan korostaa tutkimuksellista näkökulmaa, puhutaan tietokonenäöstä, (*computer vision*). Tälle suunnalle läheisiä tutkimusaloja ovat tekoälymenetelmät ja biologia (biologinen näköjärjestelmä mallina). Hahmontunnistus, joka on sukua koneoppimiselle, on tärkeä osa kuva-analyysiä ja kuvan ymmärtämistä.

Konenäköä hyödynnetään useilla teollisuuden ja arkielämän alueilla. Varhaisimpia sovelluskohteita 1980-luvulta alkaen ovat teollisuus- ja maataloustuotteiden visuaalinen laaduntarkastus, teollisuusrobottien ohjaus sekä satelliittikuvien tulkinta. Jokaiselle kansalaiselle tuttu kehitys viime vuosilta on automaattisen kasvojentunnistuksen (Li ja Jain, 2011) käyttöönotto. Rajavalvonnan passintarkastusautomaatit ovat näkyvä viranomaissovellus, mutta myös sosiaalisen median sovellukset ja digitaaliset kamerat hyödyntävät samaa teknologiaa.

2.2.3 Luonnollinen kieli ja kognitio

Luonnollisen kielen käsittely (*Natural Language Processing*, *NLP*) tarkoittaa tietokoneohjelmien käyttämistä luonnollisen tekstin ja puheen analysointiin ja tuottamiseen (Manning ja Schütze, 1999). Ala sisältää seuraavat osa-alueet:

- konekääntäminen,
- automaattinen puheentunnistus,
- puhesynteesi,
- tekstintunnistus (*Optical Character Recognition*, *OCR*),
- älykäs tekstinsyöttö ja
- puheen kääntäminen.

Ala sivuaa monessa suhteessa lingvistiikkaa, jossa puolestaan voidaan erottaa seuraavat hierarkiatasot: fonetiikka, fonologia, morfologia, syntaksi, semantiikka ja pragmatiikka. Luonnollisen kielen osa-alueet ja sen taustalla oleva lingvistiikka on jäsennetty kuviossa 4.

Luonnollisen kielen käsittelyssä on kuljettu samaa polkua kuin monilla muilla tekoälyyn luettavilla aloilla: tutkijoiden luomista säännöistä ja malleista kohti tilastolliseen dataan pohjaavia menetelmiä ja viime aikoina erityisesti neuroverkkojen hyödyntämiseen. NLP-ohjelmistojen suorituskykyä mitataan mm. NIST 2000 conversational telephone speech recognition -testillä. Puhuttujen sanojen tunnistamista mittaavassa testissä ohjelmistot ovat saavuttaneet ihmisen tason (5,9 % väärin tunnistettuja sanoja) vuonna 2016 (Xiong ym., 2016).

Käytännön sovelluksia ovat puheohjaus autoissa ja matkaviestimissä, sekä nyttemmin Sirin ja Alexan kaltaiset assistentit, sanelusta tekstiä tuottavat järjestelmät esimerkiksi lääkäreille, roskapostin suodatus, Google translatorin kaltaiset käännösohjelmat, puheen syntetisointi sekä automaattiset asiakaspalveluohjelmistot, jotka vastaavat tietyn aihepiirin kysymyksiin tai neuvontatarpeisiin. NLP-menetelmiä sovelletaan myös tiedon louhintaan ja dokumenttien automaattiseen tiivistämiseen.

Semanttinen web on alue, joka liittyy luonnollisen kielen tutkimukseen ja tekoälyyn. Semanttinen tietoverkko perustuu semantiikaltaan suoraan tekoälytutkimuksen piirissä keskeisiin

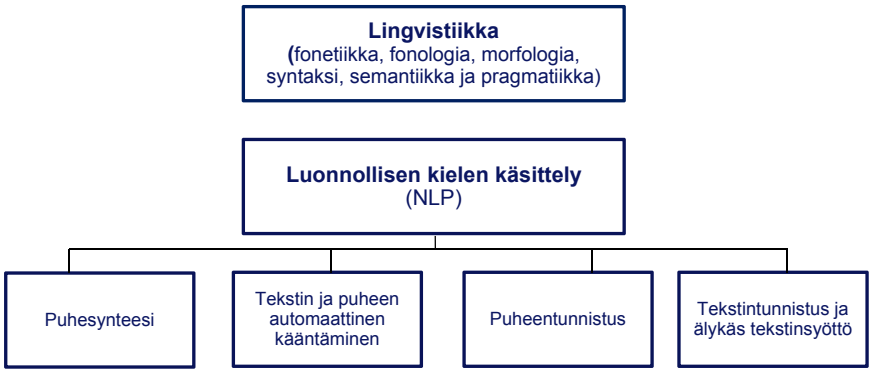
logiikkoihin ja päättelyyn (RDF, OWL yms.), ja web tarjoaa paljon kiinnostavia toimintoja, datoja, haasteita yms. tekoälyn keinoin analysoitaviksi ja ratkottavaksi (Hyvönen, 2018). Alue liittyy läheisesti luonnollisen kielen käsittelyyn, sillä webin sivuista keskeinen osa on tekstipohjaista tietoa, josta nyt louhitaan tekoälyn keinoin esiin merkitysrakenteita, semantiikkaa. Koska web on (tieto)yhteiskuntaa luultavasti eniten ja nopeimmin muuttanut teknologia sitten tietokoneen keksimisen, tämä ala on myös vaikuttavuudeltaan merkittävä. Semanttinen web voidaan nähdä osana laajempaa tekoälyn ja webin keskinäistä riippuvuutta, joihin liittyviä muita tutkimusaloja ja teknologioita ovat mm. joukkoistamisen menetelmät, luonnollisen kielen käsittely web-sovelluksissa, yksityisyyden suojaamisen menetelmät sekä webin sosiaalisten verkostojen tutkimus. Keskeisen tekoälykonferenssin (IJCAI, 2016) AI & Web Track antaa käsityksen tutkimusaiheista[5].

Kognitio on luonnollisen tai keinotekoisen systeemin kyky ennakoida tarve toimille (*action*) ja arvioida näiden tulos tai vaikutus ennakolta (Vernon, 2014). Kognitiotiede (*cognitive science*) tutkii mieltä ja sen prosesseja, ja sen perusajatus voidaan kiteyttää seuraavasti: ”ajattelua voidaan parhaiten ymmärtää mielen (käytössä olevien tiedon ja tietämyksen) esittämisen rakenteiden ja niillä operoivien tiedonkäsittely prosessien avulla” (Thagard, 2008). Kognitiotiede nähdään alana, joka yhdistää *filosofiaa, psykologiaa, tekoälyn tutkimusta, neurotieteitä, lingvistiikkaa ja antropologiaa*. Kognitiotieteen tutkimuskohteita ja -kysymyksiä ovat

- ajattelu, havainnointi ja toiminta,
- tietämys ja kielenymmärtäminen (*cognitive linquistics*, *psycholinquistics*),
- oppiminen ja kehitys, kuinka yksilö oppii ja hankkii ymmärrystä ja tietoa,
- muistin toiminta,
- tekoäly (keinona ymmärtää ihmisen kognitiota),
- huomion kohdistaminen, kuinka yksilö poimii aistien tuottamasta datasta olennaisen ja keskittää huomionsa siihen ja
- tietoisuus.

Kognitiotiede on jakautunut kolmeen koulukuntaan (*classical*, *connectionist* ja *embodied cognitive science*), minkä vuoksi on jopa esitetty kysymys, pitäisikö puhua kognitiotieteistä monikossa (Dawson, 2013). Vaikka koulukuntien erot eivät suoranaisesti kuulu tämän selvityksen piiriin, ne esitellään lyhyesti liitteessä 1.

Kuvio 4 Luonnollisen kielen käsittely ja lingvistiikka



5 http://ijcai-16.org/index.php/welcome/view/web_track

2.2.4 Vuorovaikutus ihmisen kanssa

Vuorovaikutus ihmisen kanssa on tärkeässä roolissa myös tekoälyteknologioita käyttävien järjestelmien hyödyntämisessä. Osa järjestelmistä, esimerkiksi suosittelujärjestelmät, palvelurobotit, chat-botit ja ns. henkilökohtaiset assistentit on nimenomaan rakennettu palvelemaan ihmistä ja toimimaan vuorovaikutuksessa hänen kanssaan. Myös päätöksenteon tukena toimiville järjestelmille on olennaista kyetä tarjoamaan tietoa, neuvoja ja suosituksia käyttäjille heille kussakin tilanteessa sopivalla tavalla.

Ihmisen ja tietokoneen vuorovaikutuksen tutkimus (*Human Computer Interaction*, *HCI* tai *CHI*) on tutkimusala, joka tutkii ja pyrkii selvittämään ihmisen (käyttäjän) vuorovaikutusta tietokoneen kanssa ja tietokoneen kautta. HCI-tutkimuksen taustatieteitä ovat ergonomia, psykologia ja teknologian tutkimus sekä tietojenkäsittelytiede. Ala voidaan jakaa tekoälytutkimuksen kannalta kiinnostaviin osa-alueisiin seuraavasti (Sharples, 1996):

- vuorovaikutuksen mallintaminen,
- vuorovaikutuksen suunnittelun ohjeistus,
- tietokonejärjestelmien käytettävyyden ja käyttäjäkokemuksen tutkimus ja vertailumenetelmät (*usability and user experience*),
- tietokoneiden ja ihmisen vuorovaikutuksen vaikutusten tutkiminen sekä yksilön että organisaatioiden näkökulmasta.

Uudempana tutkimusteemana tekoälyn ja HCI:n leikkausalueelle on noussut tunnetekoäly (*artificial emotional intelligence*, *affective computing*), joka tutkii järjestelmiä, jotka voivat tunnistaa, tulkita, käsitellä ja simuloida ihmisen tunnetiloja (Picard, 1995). Viime aikoina tunnetilojen tunnistamisessa on edistytty mm. kasvojen ilmeiden tulkinnan (konenäkö), äänensävyn (NLP) ja fysiologisten mittausten avulla. Toisaalta myös ihmisten asenteita pystytään analysoimaan esimerkiksi sosiaalisen median käyttäytymisen perusteella ja hyödyntämään mainonnassa ja poliittisessa vaikuttamisessa, vrt. Cambridge Analytica ja USA:n presidentinvaalit 2016.

Toisen tyyppisestä vuorovaikutuksesta ihmisen ja tekoälyä hyödyntävän järjestelmän välillä on kyse esimerkiksi autonomisten autojen ja ihmisten kanssa työskentelevien robottien (*co-robots*) tapauksessa (Sheridan, 2016). Tällöin ensimmäinen vaatimus on, että autonominen järjestelmä, esimerkiksi itse ajava auto, ei saa vahingoittaa lähelleen sattuvia ihmisiä. Järjestelmän on siis kyettävä havainnoimaan ja ennakoimaan ihmisten joskus yllättäviäkin toimia.

2.2.5 Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus

Digitaidot työelämässä perustuvat henkilöiden koulutukseen, kokemukseen ja asenteisiin. Useat tahot, mm. EU (DESI, 2018) ja sveitsiläinen kauppakorkeakoulu IMD (IMD, 2018), arvioivat kansakuntien digitaitoja ja tiivistävät nämä vertailuluvuiksi ja indekseiksi. Käytettyjä mittareita ja niihin liittyviä tutkimusmenetelmiä ovat

- koulutuksen tulosten vertailu, tunnetuimpana näistä PISA-arviointi,
- koulutuspanostusten mittaaminen, esimerkiksi osuus bkt:sta,
- koulutuksen formaalien tulosten mittaaminen, esim. korkeakoulututkinnon haltijoiden osuus aikuisista ja teknisen tai matemaattisen koulutuksen saaneiden osuus,
- työvoiman tilastointi ja analysointi, esimerkiksi tilastot tietokoneen käytöstä työssä, ohjelmointi- ja muissa ICT-tehtävissä olevien osuus sekä työuran aikana tapahtuva lisä- ja täydennyskoulutus.

Suomalaisten digitaitoja pidetään kansainvälisesti vertaillen hyvinä. IMD:n selvityksessä, jossa on mukana 63 maata, Suomi on 4. digitalisaation kokonaisuudessa. Arvio koostuu yhdeksästä osa-alueesta, joista kolme liittyy digitaitoihin työelämässä (*talent*, *training and education* & *adaptive attitudes*), ja näissä Suomen sijoitus on vastaavasti 10., 8. ja 3. Talenttien alakohdassa digitaaliset taidot Suomi on 4. Vastaavasti EU:n DESI-raportissa Suomi on sijoilla 1–6 digitaitoja työelämässä mittaavissa indikaattoreissa.

Laskennallinen luovuus (*creative computing)* ja ongelmanratkaisu vievät tekoälyä nimensä mukaisesti ennen näkemättömälle ja siten ennakoimattomalle alueelle. Samalla tämä tarkoittaa, että järjestelmälle (koneelle) annetaan vastuuta (Colton ym., 2015). Tyypillisiä esimerkkejä laskennallisesta luovuudesta ovat ohjelmien tuottamat runot, tarinat, sävellykset ja kuvat tai kollaasit. Ohjelmat kykenevät myös kirjoittamaan kouluaineita ja koevastauksia, urheilu- ja vaaliuutisia sekä syntetisoimaan mahdollisesti hyödyllisiä yhdisteitä lääke- ja kemianteollisuudelle (Smalley, 2017). Laskennallisen luovuuden katsotaan olleen merkittävässä roolissa kun AlphaGo Zero -ohjelma kehittyi nopeasti go-pelin mestariksi, mutta tämä voidaan laskea myös neuroverkkolaskennan voitoksi.

2.2.6 Koneoppiminen

Koneoppiminen (*machine learning*) on tietokonetekniikan osa-alue, jossa yleensä käytetään tilastotieteen menetelmiä, jotka antavat tietokoneille kyvyn oppia datasta (s.o. parantaa suorituskykyään tietyn tehtävän suorittamisessa) ilman eksplisiittistä ohjelmointia. Koneoppimisen menetelmiä käytetään ennustamaan ja luokittelemaan ilmiöiden tai toimien lopputulemia (Bishop, 2006). Nämä menetelmät eivät edellytä ilmiön tuntemista ja kuvaamista matemaattisella mallilla kuten esimerkiksi perinteisissä fysiikan malleissa. Koneoppiminen sopii siis tilanteisiin, joissa ilmiötä ei ymmärretä tai sen mallintaminen nähdään liian työlääksi, mutta ilmiöstä on saatavilla riittävästi dataa, jotta käytettävä menetelmä voidaan opettaa. Monet koneoppimisen tekniikat ovat läheistä sukua hahmontunnistukselle (*pattern recognition*).

Koneoppimisessa järjestelmää opetetaan ensin antamalla sille opetusdataa, jonka avulla algoritmi oppii eli käytännössä säätää algoritmin sisäiset parametrit niin, että se kykenee luokittelemaan opetusdatassa olevat tapaukset (näytteet) mahdollisimman oikein tai tekemään niiden perusteella mahdollisimman hyvän ennusteen halutusta suureesta. Seuraavassa vaiheessa eli käytössä algoritmille annetaan tuntematon näyte, ja algoritmi luokittelee sen tai tekee ennusteen halutun suureen arvosta.

Koneoppimista voidaan pitää alueena, joka yhdistää tekoälyn, tietojenkäsittelytieteen (*computer science*) ja data-analytiikan menetelmiä ja tutkimustraditiota. Sille rinnakkainen tutkimusalue on laskennallinen tilastotiede. Tätä jäsentelyä on hahmotettu kuviossa 5.

Esimerkki koneoppimisesta

Toisen asteen oppilaitoksen ongelmana ovat kevätlukukauden aikana tapahtuvat ensimmäisen vuoden opiskelijoiden monet opintojen keskeytykset. Koneoppimisen avulla pyritään tunnistamaan suuren keskeytysriskin oppilaat ja kohdistamaan tukitoimet heihin. Syöttötietona ovat opiskelijan poissaolot syyslukukaudella, peruskoulun päättötodistuksen keskiarvo, opintojen alussa tehdyn motivaatiokyselyn tulokset ja vanhempien koulutustaso. Opetusdatana ovat kolmen edellisen vuoden opiskelijoiden vastaavat tiedot ja toteutuneet keskeyttämiset.

Koneoppimisen menetelmät voidaan jakaa ohjattuun ja ohjaamattomaan oppimiseen. Ohjatussa opetuksessa algoritmille kerrotaan opetusvaiheessa kunkin tapauksen (näytteen) oikea tulos, luokka tai suureen arvo.

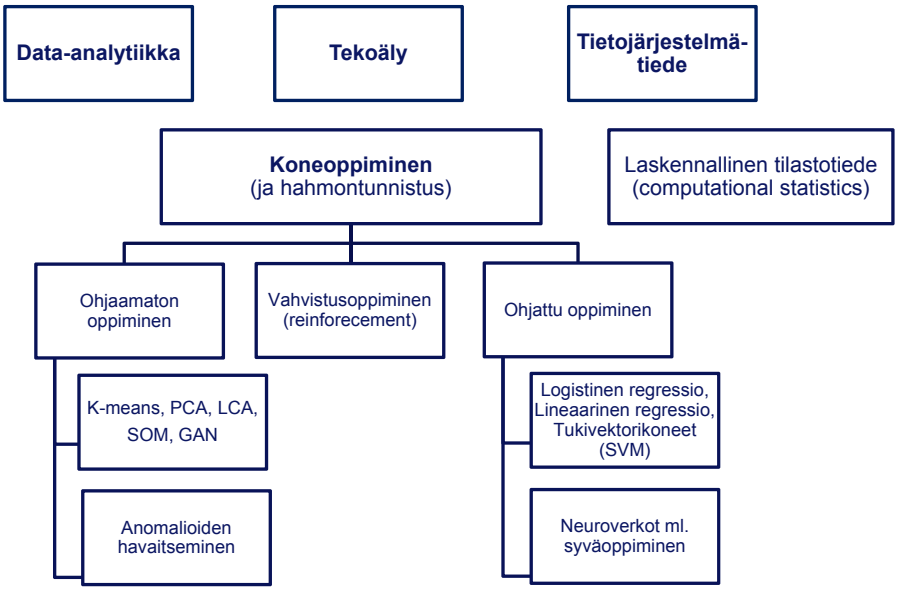
Ohjatun opetuksen tunnetuimpia menetelmiä ovat lineaarinen ja logistinen regressio, päätöspuut, neuroverkot, k-lähimmän-naapurin -menetelmä, lineaarinen erotteluanalyysi (*Linear Discriminant Analysis*, *LDA*) ja tukivektorikone (*Support Vector Machine*, *SVM*).

Ohjaamattomassa oppimisessa koneoppimisalgoritmi pyrkii itse ryhmittelemään eli klusteroimaan tapaukset eri luokkiin. Ohjaamattoman oppimisen menetelmiä ovat ryhmittely eli klusterointi, jota voidaan toteuttaa esimerkiksi k-means-algoritmilla, poikkeamien havaitseminen (*anomaly detection*), neuroverkko-menetelmät, joihin myös Kohosen kartta (*self-organizing map*) luetaan, sekä latenttien muuttujien mallit, joista pääkomponenttianalyysi lienee tunnetuin. Viimeisimpänä kehityksenä ohjaamattomassa opetuksesta voidaan nähdä ns. GAN-verkkojen (*generative adversarial network*) kehittyminen. GAN-verkkojen avulla koneille luodaan ihmisten mielikuvituksen kaltaisia, rajallisia kyvykkyyksiä (Giles, 2018).

Vahvistava oppiminen (*reinforcement learning*) sijoittuu ohjatun ja ohjaamattoman opetuksen väliin ja sopii tilanteisiin, joissa ohjauspalautetta on saatavissa rajoitetusti (Otterlo ja Wiering, 2012). Vahvistavassa oppimisessa keskeisiä käsitteitä ovat agentti (oppija), ympäristö, tilanteen tulkinta ja palkkio. Toiminta ja oppiminen tapahtuvat ajassa, jolloin nykyinen tilanne perustuu aiempaan tilaan ja siinä tehtyihin ratkaisuihin. Vastaavasti tuleva tila riippuu ympäristöistä ja agentin ratkaisuista. Oppiminen kuvataan markovilaisena päätösprosessina *(Markov Decision Process*). Menetelmän juuret ovat käyttäytymispsykologiassa ja päätöksenteon teoriassa.

Niin sanottujen syvien neuroverkkojen (*deep neural networks*) (LeCun ym., 2015) avulla on saavutettu viime vuosina erittäin merkittäviä ja näyttäviä tuloksia mm. kuvantunnistuksen, puheentunnistuksen ja kielen kääntämisen alueilla. Myös useat verkkokauppojen käyttämät ns. suosittelukoneet perustuvat neuroverkkoihin.

Kuvio 5 Koneoppiminen



Neuroverkkojen menestys on johtanut siihen, että keskustelussa usein vedetään yhtäläisyysmerkit tekoälyn ja neuroverkkoihin perustuvan koneoppimisen välille. Tämä on liian kapea näkökulma. Koneoppimisen rooli voidaan nähdä toisin: varsinainen tekoäly on koneoppimisen tulosten päälle rakentuva päätöksentekokerros (Malone, 2018). Neuroverkkoja kohtaan on esitetty myös kovaa kritiikkiä, jonka mukaan niiden heikkouksia ovat läpinäkymättömyys, ne ovat ns. mustia laatikoita, joiden tekemän päättelyn tai laskennan perusteita on lähes mahdoton tarkistaa jälkikäteen; ne käyttävät paljon laskentatehoa ja energiaa; ja ne ovat hauraita (*brittle*), ts. ovat herkkiä syöttötiedon muutoksille ja virheille. Edellä mainitut ongelmat ovat oikeustieteen näkökulmasta kriittisiä ja voidaan nähdä esteiksi teknologian hyödyntämisessä viranomais- ja tuomiovallan käytössä.

Koneoppimisen sovellus- ja hyödyntämiskohteita ovat mm. kasvojentunnistus, kuvahaut, ja autonomiset ajoneuvot, erityisesti konenäköön perustuva ohjaus.

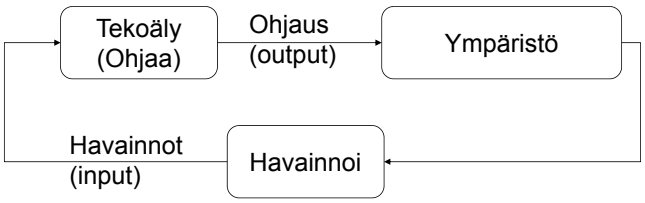
2.2.7 Järjestelmätaso ja systeemivaikutukset

Järjestelmätaso ja systeemivaikutukset ymmärretään tässä näkökulmaksi, jossa tekoälyteknologioita ja niiden käyttöä tarkastellaan laajemman järjestelmän kokonaisuuden näkökulmasta (Sage, 1992), ikään kuin ylhäältä päin. Läheisiä insinööritieteen aloja ovat

- säätö- ja systeemitekniikka,
- prosessinohjaustekniikka,
- systeemidynamiikka,
- elinkaarimallinnus ja
- tuotantotalous.

Systeemiteknistä (*systems engineering*) lähestymistapaa yleisen tekoälyn haasteisiin (*General AI*, *AGI*) on esitetty tuoreessa artikkelissa (Alpcan ym., 2017). Yleinen tekoäly nähdään mahdolliseksi, kun yhdistetään suosittuja datapohjaisia (*connectionist*) menetelmiä, kuten syväoppivia neuroverkkoja, tilastollista hahmontunnistusta ja klassista symbolista tekoälyä systeemiteorian ohjauksessa. Kuvio 6 tiivistää kontrolliteorian mukaisen näkemyksen tekoälyn rooliin.

Kuvio 6 Kontrolliteorian mukainen yksinkertaistettu näkemys tekoälyyn



Lähde: Alpcan ym. (2017) mukaillen.

Keskeisiä tutkimuskysymyksiä järjestelmätason ja systeemivaikutusten näkökulmasta ovat

- datapohjaisten ja symbolisten tekoälymenetelmien yhdistäminen,
- toisten tieteenalojen ja menetelmien hyödyntäminen tekoälyn kehittämisessä: optimointi, informaatioteoria, peliteoria ja ohjausteoria, koska tekoälytutkimuksen kohtaamia ongelmia on kohdattu ja myös ratkaistu aiemmin muilla aloilla,

– biologisten järjestelmien ja luonnollisen älykkyyden parempi ymmärtäminen, koska se voi johtaa edistysaskeliin yleisen tekoälyn suhteen.

Toinen näkökulma on tekoälymenetelmien käyttäminen systeemisuunnittelun välineenä tai apuna. Insinööritieteen aloja, joilla tekoälyä käytetään suunnittelun apuna ovat elektroniikkapiirien suunnittelu, ohjelmointi, rakennus- ja infrasuunnittelu, konetekniikka, kemiantekniikka ja biolääketiede. Erityisesti kahteen viimeksi mainittuun alaan liittyvissä kemiallisten yhdisteiden simuloinnissa ja seulonnassa tekoälymenetelmät ovat osoittautuneet hyödyllisiksi (Smalley, 2017). Edelleen systeemiteknistä näkökulmaa voidaan käyttää myös itse tekoälyjärjestelmien suunnittelussa (Erman ja Lessner, 1977).

2.2.8 Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit

Tekoälymenetelmien vaatimia ympäristöjä voi yksinkertaistaen tarkastella pinona, jossa alimpana on laskentayksikkö eli prosessori, sen yläpuolella käyttöjärjestelmä ja ohjelmointikieli, ylimpänä itse varsinainen tekoälysovellusohjelma.

Tarkastellaan aluksi tekoälylaskennassa käytettyjen prosessorien kehitystä. Tekoälysovellusten ja tutkimuksen ohjelmistototeutuksia on 1980-90-lukujen vaihteen LISP-koneiden (esim. Symbolics) jälkeen ajettu viime vuosiin asti yleiskäyttöisillä tietokoneilla. Tilanne kuitenkin muuttui, kun koneoppimisen, erityisesti syvien neuroverkkojen, sovellukset alkoivat vaatia erittäin suurta laskentatehoa. 2010-luvulla ns. graafiset prosessorit (*Graphical Processing Unit*, *GPU*), joita käytetään esim. pelikoneissa, osoittautuivat tehokkaaksi neuroverkkolaskennassa. Neuromorfisessa laskennassa pyritään analogia- ja digitaalitekniikkaa yhdistämällä rakentamaan prosessoreita, jotka matkivat ihmisen aivojen hermosolujen toimintaa paremmin kuin tavalliset digitaaliset prosessorit. EU:n rahoittama Human Brain

USA:n ja Kiinan kilpa – missä Eurooppa?

Kiina on julistanut tekoälyn kansalliseksi prioriteetiksi presidentti Xin arvovallalla ja näin haastaa Yhdysvaltain johdon alalla. Maa aikoo sijoittaa 150 miljardia dollaria tutkimukseen ja kehitykseen ja nousta johtavaksi tekoälyn jättiläiseksi vuoteen 2030 mennessä (Metz, 2018). Tekoälyn nostamisessa kansalliseksi prioriteetiksi Kiina tavallaan seuraa Obaman kauden Yhdysvaltoja, joka ei presidentti Trumpin aikana ole ottanut kansallisella tasolla kantaa tekoälyyn.

Tekoälyn kenttää hallitsevat vielä amerikkalaiset yritykset, mutta kiinalaisten sijoitukset kasvuyritykseen ulkomailla ja kotimaassa ohittivat Yhdysvaltain vastaavat 2017 ja myös patenttien ja julkaisujen määrässä johtoasema on vaihtunut tai vaihtumassa.

EU ja sen jäsenmaat ovat syystäkin huolissaan tilanteesta. Eurooppalaiset yritykset eivät ole onnistuneet pääsemään tekoälyalustojen toimittajiksi tai ekosysteemien keskeisiksi toimijoiksi, vaikka täällä on suhteellisen paljon start-up -yrityksiä. Uhkana on, että Eurooppa jää sivustakatsojaksi USA:n ja Kiina jakaessa tekoälymarkkinan. EU:n komissio on reagoinut tähän julkistamalla huhtikuussa 2018 virallisen kommunikean Artificial Intelligence for Europe, jossa asetetaan tavoitteet teknologiakehitykselle, yhteiskunnan muutosvalmiuden parantamiselle sekä eettisen ja juridisen kehikon luomiselle.

-tutkimusprojekti lähestyy aihetta sekä biologian että tekniikan näkökulmasta. Neuromorfiset prosessorit ovat vielä tutkimusasteella.

Viime vuosien aikana sovellusten koko ja käytettävät datamäärät ovat kasvaneet erittäin voimakkaasti, joten oli siirryttävä yksittäisistä tietokoneista kohti pilvilaskentaa ja ns. serverifarmien käyttöä. Samalla liiketoimintamallit ovat muuttuneet.

Ala voidaan jaotella tutkimuksen näkökulmasta seuraaviin osa-alueisiin:

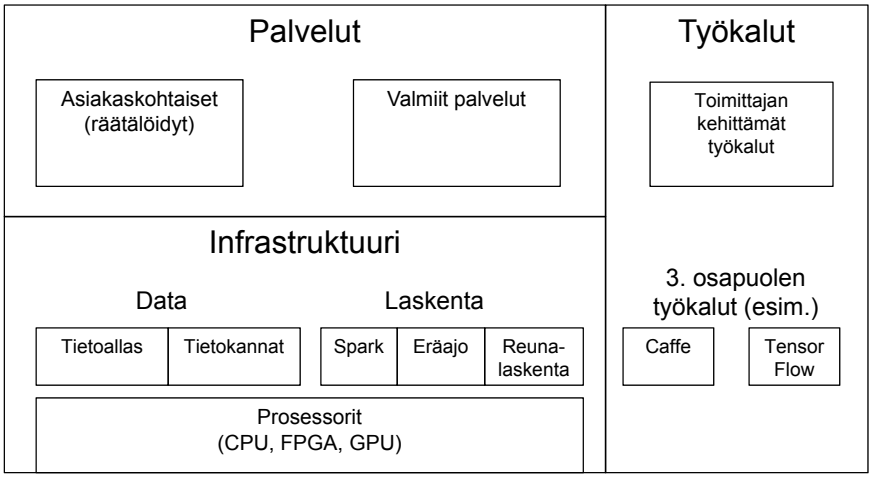
- prosessoritekniikka (ml. uudet neurolaskentaan suunnitellut arkkitehtuurit),
- tietokonejärjestelmät ja arkkitehtuurit (erityisesti massadatan käsittelyä palvelevat pilvilaskennan järjestelmät),
- tietoliikenne ja verkkoteknologiat,
- ohjelmistoteknologia ja -arkkitehtuurit, hajautettu laskenta, työkalut,
- tietosuoja, anonymisointi, kryptografia,
- liiketoimintamallien tutkimus ja kehittäminen.

Koska monet koneoppimisen menetelmät vaativat massiivisia opetusaineistoja, tarvitaan ns. Big Data -teknologioiden käyttöä datan tallentamiseen ja piirteiden laskentaan (*feature extraction* / *engineering*). Tämä on yksi syy sille, että tekoälyn hyödyntäminen on erittäin pääomavaltainen ala. Toinen syy on monesti vielä opetusaineiston osittain manuaalinen (ihmistyöllä) tehtävä labelointi, mutta siitä tietenkin pyritään pääsemään eroon. Usein algoritmien hyvyys perustuu juurikin massiivisten opetusjoukkojen käyttöön. (”*It’s not who has the best algorithm that wins, It’s who has the most data*”, Andrew Ng).

Viimeisten vuosien aikana tilanne on kehittynyt suuntaan, jossa laskennan suorittavat prosessorit sijaitsevat usein muualla, ns. pilvessä, ja koko laskenta suoritetaan palveluntarjoajan, esimerkiksi Amazonin *alustalla* ja se ostetaan *palveluna* (ks. myös kuvio 7). Eri toimijat muodostavat verkoston, joka yhdessä toteuttaa halutun tekoälylaskennan tai sovelluksen. Tällä hetkellä koneoppimiseen pohjaavat teknologiat ovat pääroolissa myös laskentaympäristöjen, alustojen ja palveluiden kehittymissä.

Tekoälyyn pohjautuvaa ratkaisua tarvitseva asiakas voi siis käyttää palveluntarjoajan alustaa ja työkaluja, joilla ensin kehitetään tarvittava sovellus ja myöhemmin tuotantokäytössä

Kuvio 7 Kaupallinen tekoälysovellusten kehitys- ja sovellusympäristö



ajetaan sitä. Palveluntarjoajan alustan kautta asiakas saa käyttöönsä laskentaympäristön, jonka osana ovat yllä kuvatut tekoälyprosessorit sekä suurten datamäärien tallentamiseen ja siirtoon tarvittavaa kapasiteettia. Palveluntarjoajan myyntivaltteja ovat skaalautuvuus, koetellut ratkaisut, kiinteiden investointien pienuus ja hyvät kehitystyökalut, jotka tuovat myös etua kehityksen ja käyttöönoton nopeudessa. Heikkouksina voi nähdä datan hallintaan liittyvät kysymykset (yksityisyyden suoja, suojattava viranomaistieto, liikesalaisuuksien suoja) ja joissain tapauksissa räätälöintimahdollisuuksien heikkoudet.

Merkittävimpiä tekoälylaskennan palveluntarjoajia ja samalla ekosysteemien päätoimijoita ovat Microsoft (Azure-tuoteperhe), IBM (Watson markkinointinimenä), Google (Brain ja DeepMind), Amazon (AWS) sekä kiinalaiset Alibaba, Baidu ja nyttemmin Tencent. Näistä Microsoft ja IBM ovat perinteisempiä IT-toimijoita, kun taas Amazon ja Alibaba ovat verkkokaupan jättiläisiä, jotka ovat onnistuneesti luoneet synergiaa verkkokaupasta pilvipalveluihin ja edelleen tekoälyalustoihin. Googlen ja Baidun tausta on hakukonemainonnassa, joka on koneoppimisen tärkeä hyödyntäjä.

Tekoälylaskennan ekosysteemeihin liittyvät prosessorien valmistajat, osa suoraan, osa väljemmin. Merkittävimpiä näistä ovat NVIDIA ja Intel, mutta myös AMD:llä, Googlella ja Applella on näkyviä tuloksia alalla (Googlen Tensor Processing Unit TPU ja Applen A11-koneoppimisprosessori puhelimiin). Kiinan strateginen tekoälypanostus kattaa myös AI-prosessorit, ja maan saavutukset super-tietokoneissa asettavat odotukset korkealle. Kiinalaisista AI-prosessorien kehittäjistä kiinnostavia ovat Huawei (koneoppimisprosessori puhelimeen, kuten Applella) ja Lenovon rahoittama kasvuyritys Cambricon.

Kuviossa 7 on esitetty yksinkertaistettu kaavio kaupallisesta pilvipalveluna toimivasta tekoälysovellusten kehitys- ja suoritusympäristöstä.

2.2.9 Robotiikka ja koneautomaatio – tekoälyn fyysinen ulottuvuus

Roboteilla tarkoitetaan laitteita[6], jotka kykenevät vaikuttamaan fyysiseen ympäristöönsä esimerkiksi tarttujan, käsivarren, pyörien tai jalkojen avulla (Russell ja Norvig, 2014). Robotit toimivat yleensä tietokoneohjelman ohjaamina. Robotilla voi olla aisteja, esimerkiksi konenäkö ja tuntoaisti.

Robotiikka on tutkimus- ja teknologia-ala, joka sivuaa mekaniikkaa, sähkö-, säätö ja automaatiotekniikkaa sekä tietokone- ja tietojärjestelmätekniikkaa. Joskus myös konenäkö liitetään robotiikkaan, vaikka se on myös itsenäinen tutkimusala. Robotiikkaa voidaan pitää koneautomaation osa-alueena, mutta usein se käsitetään omaksi tutkimus- ja teknologia-alaksi, jonka juuret ovat koneautomaatiossa.

Tutkimusalana robotiikka voidaan jakaa seuraaviin osiin (Russell ja Norvig, 2014):

– Robotin tehtäväsuunnittelu (*task planning*, *mission planning*) tutkii ja kehittää menetelmiä ja rakenteita, joilla robotti parhaiten suoriutuu tehtävistä, joita varten se on olemassa. Tehtäväsuunnittelu muodostaa robotin ylimmän tason, joka ohjaa havainnointia, liikkumisen ja liikeratojen suunnittelua ja toimilaitteita. Myös ohjelmistoarkkitehtuurien suunnittelussa huomioidaan tehtäväsuunnittelun paradigmat.

6 Ns. ohjelmistorobotteja tai ”botteja” ei tässä käsitellä, vaan rajoitutaan laitteisiin, joilla on fyysinen ulottuvuus.

– Toimilaitteiden (*effectors*) tutkimus, jossa keskitytään robotin fyysisen liikkeen tuottaviin osiin, kuten sähkö- tai pneumaattisiin moottoreihin, niveliin ja varsiin, joiden liikkeiden tuloksena syntyy robotin kinemaattinen tila (paikka ja asento) (Craig, 2005). Vapausaste (*degrees of freedom*) on keskeinen käsite, joka kuvaa robotin ja sen nivelten liikkuvuutta (kiertoja tai lineaariliikettä).

– Anturien ja havaitsemisen (*sensors and perception*) tutkimus, jonka tarkoitus on kehittää menetelmiä, joilla epävarmasta anturitiedosta voidaan estimoida mahdollisimman hyvin käsitys robotin ympäristöstä ja omasta tilasta (esim. Zhuang ja Roth, 1996). Erityisesti ollaan kiinnostuneita ympäristön esineiden, kuten työkappaleiden tai esteiden paikantamisesta (*localization*) sekä robotin omasta paikasta, nivelten asennoista (*pose*) sekä liiketilasta. Havaitsemiseen liittyy aikasarjojen käyttö, ts. nykyisen tilan estimoinnissa käytetään paitsi antureiden tuottamaa tietoa, myös edellistä estimaattia. Tehtävän luonteen takia tila-arviot ovat todennäköisyyksiin perustuvia ja tyypillisiä menetelmiä ovatkin Kalman-suodattimet ja Bayesilaiset verkot.

– Liikkeen ja liikeratojen suunnittelu on robotin toiminnan kannalta keskeinen tutkimusalue, jossa tutkitaan esimerkiksi, miten robotin koura saadaan optimaalisesti ohjattuun haluttuun pisteeseen (jos robotilla on useita niveliä, tämä voidaan tehdä äärettömän monella eri tavalla), tai miten robotti toimii käsitellessään työkappaleita (Vukobratovic ym., 2009) (*compliant motion*), esimerkiksi asettaessaan auton tuulilasin särkymättä paikalleen. Liikkeen ja liikeratojen suunnittelussa käytetään ns. kustannusfunktiota (*cost function*), jonka suhteen suunnitelma optimoidaan. Käytännön tilanteissa suunnitteluun liittyy olennaisesti epävarmuuden käsite, joka johtuu robotin dynaamisesti muuttuvasta ympäristöstä ja anturitiedon epäluotettavuudesta. Suunnittelu epävarmuuden vallitessa onkin keskeinen tutkimushaaste liikkeiden suunnittelussa.

– Suunnitellun liikeradan toteuttaminen todellisella robotilla on tutkimushaaste, jota lähestytään dynaamisen mallin ja ohjauksen (*dynamics*, *control*) pohjalta. Käytettyjä lähestymistapoja ovat säätö- ja systeemitekniikan menetelmät, kuten PID-säätäjät, reaktiivinen ohjaus, jota käytetään esimerkiksi kävelevissä roboteissa, sekä vahvistava opetus (*reinforcement learning*), jossa drone-lentolaitetta ohjaava toimintatapa (*policy*) voidaan oppia taitavalta lennättäjältä.

– Robottia ohjaavien tietokoneohjelmien järjestystä ja rakennetta kutsutaan ohjelmistoarkkitehtuuriksi. Arkkitehtuurin tehtävä on esimerkiksi päättää, missä roolissa on liikkeiden toteuttamiseksi tehty suunnitelma suhteessa reaktiiviseen ohjaukseen. Ohjelmistoarkkitehtuurien tutkimuksessa pyritään valitsemaan ja kehittämään (eri tehtäviin) optimaalinen ohjelmistoarkkitehtuuri. Vaihtoehtoja ovat esimerkiksi kerros- tai pinoarkkitehtuuri, joka sisältää reaktiivisen tason (välitön reaktio ympäristön yllättäviin muutoksiin), toimeenpanotason (suunnitelmien toteutus) ja suunnittelutason, sekä toimenpideputki-arkkitehtuuri (*pipelane*). Kontrolliarkkitehtuurin tutkimuksessa haasteena on reaktiivisen ja deliberatiivisen lähestymistavan optimaalinen yhdistäminen päämääränä suunnitellun tavoitteen saavuttaminen.

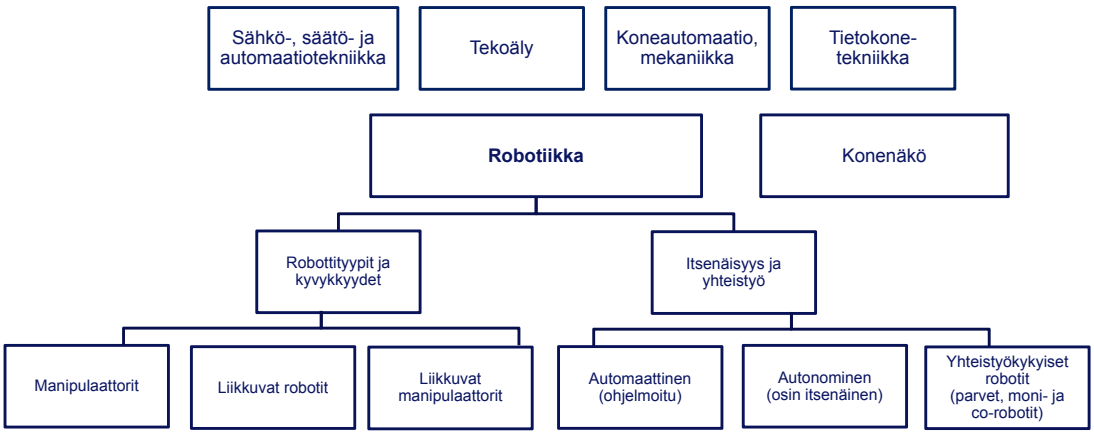
Robotteja voidaan luokitella niiden kyvykkyyksien tai käyttötarkoituksen mukaisesti. Kyvykkyyden perusteella robotit jaetaan automaattisiin ja autonomisiin. Automaattinen robotti suorittaa sille ennalta määrättyä (ohjelmoitua) tehtävää tunnetussa ympäristössä, esimerkiksi tuotantolinjalla. Robotti suorittaa tehtävänsä tarkoin ohjelman mukaisesti ja kykenee mukautumaan vain pieniin muutoksiin esimerkiksi työkappaleen paikan suhteen. Sen sijaan

autonominen robotti kykenee reagoimaan ja tekemään päätöksiä muuttuvassa ympäristössä suorittaakseen sille annetun tehtävän (Correll, 2016).

Käyttötarkoituksen mukaisesti robotit jaetaan kolmeen ryhmään: manipulaattoreihin, joita ovat paikalleen asennetut teollisuusrobotit, liikkuviin robotteihin, kuten maalla pyörillä, telaketjuilla tai jaloilla liikkuvat laitteet sekä UAV:t (*unmanned aerial vehicle*) ja drone-lentolaitteet, sekä liikkuviin manipulaattoreihin, jotka ovat edellisten yhdistelmä (Russell ja Norvig, 2014).

Robotiikkaan liittyvät muut tutkimusalat ja toisaalta robotiikan tutkimuksen ja robottien taksonomiaa on esitetty kuviossa 8.

Kuvio 8 Robotiikan kentän jäsentely



Aiemmin robotiikan tutkimus ja kehitys keskittyivät yksittäisen robotin kyvykkyyden parantamiseen, mutta jo useiden vuosien ajan on tutkittu ja kehitetty yhdessä toimivia monirobotteja (*multi-robots*), parviälyä ryhmänä toimivien autonomisten robottien tarpeisiin sekä ihmisen ja muiden toimijoiden kanssa turvallisesti toimivien robottien (*co-robots*) tai autonomisten koneiden ohjelmoimiseksi. Parviälyn näyttäviä manifestaatioita ovat ryhminä toimivat drone-lentolaitteet. Autonomisten lentolaitteiden ohjauksen haasteet ovat toimineet robotiikan, säätötekniikan ja tekoälyn tutkimuksen koetinkivenä ja testinä viime vuosina tuottaen näyttäviä tuloksia.

2.2.10 Etiikka, moraali, regulaatio ja lainsäädäntö

Insinööritieteet pyrkivät vaikuttamaan ja muuttamaan maailmaa, eivät pelkästään selittämään sitä kuten luonnontieteet. Sen vuoksi niihin liittyy aina eettisiä ja moraalisia kysymyksiä (Poel ja Royakkers, 2001). Vastaavasti lainsäädäntö ja muu sääntely (regulaatio) ovat vuorovaikutuksessa teknologian kanssa. Kun tekoälyä tarkastellaan teknologiana ja insinööritieteenä, nämä huomiot luonnollisesti koskevat sitä.

Tekoälyteknologioiden näkökulmasta voidaan tunnistaa seuraavat merkittävät etiikan ja moraalin osa-alueet.

– Moraalifilosofia, joka tutkii alan peruskäsitteitä (metaetiikka), ja normatiivisen etiikan suuntia (velvollisuus-, hyve- ja seurausetiikka).
– Soveltava etiikka, jossa pyritään etsimään vastauksia käytännön eettisiin ongelmiin soveltamalla filosofisen etiikan teorioita. Esimerkiksi sairaaloissa toimii eettisiä työryhmiä, jotka hyödyntävät soveltavan etiikan tuloksia tehdessään tapauskohtaisia päätöksiä potilaiden hoidosta.
– Teknologian etiikka, jossa tutkimuskysymyksinä ovat yhtäältä teknologian suorat vaikutukset ja toisaalta teknologian aiheuttamat välilliset vaikutukset ihmisiin ja yhteiskuntaan.
– Sodankäynnin etiikan alalla tutkitaan myös uusien teknologioiden nousun eettisiä ja moraalisia kysymyksiä (military ethics and emerging technologies) (USNWC, 2018).

Etiikan ja moraalin pohdinnan osalta tekoäly poikkeaa olennaisesti perinteisestä teknologian etiikasta, jossa eettisen ja moraalisen päätöksenteon toimijana nähdään teknologiaa suunnitteleva ihminen (Polonski, 2017). Tämän totutun lähtökohdan lisäksi joudutaan pohtimaan kahta muuta kysymystä. Ensiksi, miten opettaa etiikkaa itsenäisiä päätöksiä tekevälle tekoälylle. Kone-etiikka (*machine ethics*) tarkastelee asiaa koneen näkökulmasta (Anderson, M. ja Anderson, S., 2011). Tällöin avainkysymys on, kuinka voimme opettaa tekoälyn tekemään eettisiä valintoja. Oikeustieteen perspektiivistä kysymys on siitä, miten kone voisi oppia tekemään juridista harkintaa, esimerkiksi ottamaan huomioon suhteellisuusperiaatteen tehdessään päätöstä. Populaari esimerkki tästä on kysymys, mitä itseajava auto tekee tilanteessa, jossa onnettomuus on väistämätön. Toinen kysymys on, muodostuuko tekoälyagentille jossain vaiheessa omaa moraalista arvoa ja oikeuksia (vrt. eläinten oikeudet).

Kymmenen kysymystä tekoälyn etiikasta ja vaikutuksista

Maailman talousfoorumi (WEF, 2016) on kiteyttänyt tekoälyyn liittyvät keskeiset kysymykset seuraavasti

1. Työttömyys: Viekö tekoäly työpaikat?
2. Epätasa-arvo: Johtaako tekoälyn käyttö varallisuuden entistä suurempaan keskittymiseen?
3. Inhimillisyys: Vaikuttaako tekoäly ja robotit ihmisten käyttäytymiseen ja kanssakäymiseen?
4. Kuinka suojautua tekoälyn virheiltä ja virhetoiminnoilta?
5. Tekoälyn puolueellisuus: Oppiiko kone ennakkoluuloiseksi?
6. Kuinka suojata tekoälyjärjestelmiä pahantahtoisilta toimijoilta?
7. Pullon henki: Voiko tekoälyllä olla ei toivottuja sivuvaikutuksia?
8. Singulariteetti: Miten suojautua mahdollisesti vallanhaluiselta tekoälyltä?
9. Tekoälyn ja robottien oikeudet: Jos koneille kehittyy tietoisuus, kuuluuko niille oikeuksia kuten eläimille tai ihmiselle?

Listan alkupään kysymykset ovat akuutteja jo nyt, kapean tekoälyn aikana. Kaksi tai kolme viimeistä kysymystä tulevat ajankohtaisiksi, jos tutkimuksessa saavutetaan laadullista kehitystä yleisen (vahvan) tai jopa supertekoälyn suuntaan. Osa kysymyksistä liittyy taloustieteisiin (1, 2), yhteiskuntafilosofiaan, politiikkaan ja yhteiskunnallisiin arvoihin (2, 3, 5), osa taas on luonteeltaan teknisiä (4) ja liittyy kyberturvallisuuteen (6), kun taas kysymykset 7–9 ovat luonteeltaan moraalifilosofisia.

Algoritmeihin ja digitalisaatioon pohjautuvan päätöksenteon tulo voidaan nähdä jopa julkishallinnon ja oikeudenkäytön legitimiteetin lähteen paradigmamuutoksena (Suksi, 2017). Tämä vertautuu 1700-luvun muutokseen, jossa jumalasta kuninkaan kautta hallintoon delegoituvan legitimiteetin sijaan nousi yhteisen kielen, kansan ja kansallisvaltion tuottama legitimiteetti. Jo nyt monet hallinnolliset ja oikeudelliset päätökset tekee algoritmi tai tietokoneohjelma automaattisesti ja välittömästi. Algoritminen päätöksenteko sopii nyt tapauksiin, joissa toimitaan suurten tietomassojen kanssa selvien sääntöjen pohjalta, joista algoritmit johdetaan. Tällaisia ovat etuisuuspäätökset, verotus ja ylinopeussakot. Suksi (2017) ennakoi algoritmisen automaattisen päätöksenteon laajenevan monimutkaisempiin tapauksiin, jos tekoälyteknologioissa tapahtuu edistystä yleisen tekoälyn suuntaan.

Tekoälyyn liittyviä eettisiä kysymyksiä pohtimaan on perustettu Future of Life -insitituutti Oxfordin yliopiston yhteyteen (Future of Life, 2017). Keskeinen tutkimuskysymys heidän näkökulmastaan on, poikkeaako tekoäly eettisestä näkökulmasta muista (uusista) teknologioista.

Kuten todettiin, lainsäädäntö ja muu sääntely ovat vuorovaikutuksessa teknologian kehittämisen ja soveltamisen kanssa (OECD, 2011). Sääntely voi edistää tietyn teknologian kehitystä ja kaupallista menestystä, kuten aikanaan NMT ja GSM edistivät matkapuhelinten käyttöönottoa ja niihin liittyvää liiketoimintaa Suomessa ja Ruotsissa. Sääntely voi myös hidastaa kehitystä: autonomiset autot eivät yleisty, jos niiden kokeileminen ja käyttö on kiellettyä tai sisältää suuren taloudellisen riskin.

Sääntelyn ja standardoinnin suhdetta teknologian kehitykseen käsitellään innovaatiotutkimuksen piirissä (Blind, 2013). Tekoälyn standardointi on käynnistynyt standardoimisjärjestöjen (ISO, ITU ja ANSI) piirissä. Teknologisen sääntelyn (esimerkiksi standardit) lisäksi sääntely voidaan jakaa kolmeen alalajiin: taloudelliseen sääntelyyn, jonka tarkoitus on markkinoiden toiminnan tehostaminen; sosiaaliseen sääntelyyn, joka tähtää ympäristön, yhteiskunnan ja sen jäsenten suojaamiseen esimerkiksi terveysriskeiltä; ja hallinnolliseen sääntelyyn, joka ohjaa julkisen ja yksityisen sektorin käytäntöjä ja toimia (OECD, 2011).

Tekoälyteknologioiden ja lainsäädännön keskinäisen vaikutuksen kannalta kiinnostavia tutkimus- ja osaamisalueita ovat seuraavat:

- datan hallintaa ja omistajuutta koskeva lainsäädäntö
- yksityisyyden suojaa koskeva lainsäädäntö, esimerkiksi EU:n yleinen tietoturva-asetus GDPR
- kansalaisten yhdenmukaista kohtelua koskeva lainsäädäntö (vrt. koneoppimisalgoritmien tekemät päätökset, ja niihin sisältyvä bias)
- tekoälyteknologioiden käytön seurauksia koskevat vastuukysymykset (juridinen vastuu, korvausvelvollisuudet)
- kilpailulainsäädäntö, ml. määräävän markkina-aseman käyttöä koskeva lainsäädäntö
- teknologian ja kansallisen turvallisuuden suhdetta koskeva lainsäädäntö (datan keruu, yhdistäminen ja analysointi)
- kansainväliset sopimukset tekoälyteknologioiden käyttämisestä asejärjestelmissä ja sodankäynnissä (automaattiset ja autonomiset asejärjestelmät, robottisotilaat).

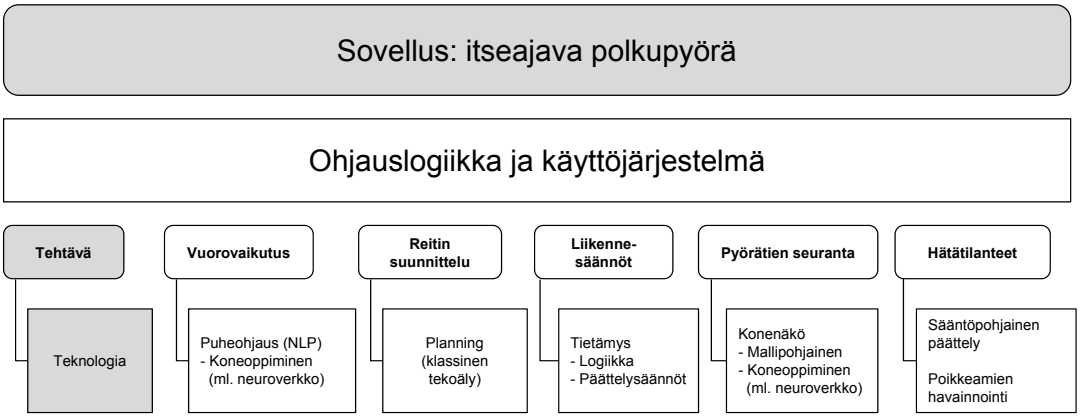
Euroopan komission tuore tekoälyteknologioita koskeva julistus edellyttää soveltuvan eettisen (ohjeistuksen) ja lainsäädäntökehikon luomista tekoälyteknologioiden soveltamiseen. Julistus edellyttää lainsäädännön olevan sopusoinnussa EU:n arvojen ja perusoikeuslinjausten kanssa. Lisäksi lainsäädännön tulee määritellä vastuukysymykset ja sitä luotaessa tulee ottaa kantaa tuleviin haasteisiin ja määritellä eettinen ohjeisto.

Ei ole vielä mitenkään selvää, miten lainsäädäntövalta jakautuu EU:n ja jäsenvaltioiden välillä, kun tekoälyä aletaan säännellä. Tällä hetkellähän sääntely on vielä sangen puutteellista, oli sitten kyse EU:sta tai jäsenvaltiosta.

2.3 Esimerkki tekoälyteknologioiden käytöstä

Seuraavassa on hahmoteltu kuvitteellisen esimerkin avulla tekoälyteknologioiden käyttöä ja sitä, että suhteellisen yksinkertainenkin sovellus hyödyntää useita ns. tekoälyteknologioita sekä perinteisempää systeemi- ja ohjelmistoteknologiaa (kuvio 9). Esimerkin järjestelmä on itseajava sähköpyörä, jonka ohjausjärjestelmä sisältää seuraavat toiminnot ja niiden toteuttamiseen tarvittavat teknologiat (suluissa): käyttäjän ja koneen välinen vuorovaikutus (puheohjaus, koneoppiminen), reitinsuunnittelu (*planning*, klassinen tekoäly), liikennesääntöjen huomioiminen (logiikka ja päättelysäännöt), pyörätien seuranta (mallipohjainen ja neuroverkkopohjainen konenäkö), hätätilanteiden hallinta (sääntöpohjainen päättely ja poikkeamien havaitseminen).

Kuvio 9 Kuvitteellinen esimerkki, jossa hyödynnetään useita tekoälyteknologioita



3 HUIPPUOSAAMISEN KARTOITUS

3.1 Johdanto osaamiskartoitukseen

Tekoäly ei ole yksi teknologia, vaan joukko menetelmiä, teknologioita, sovelluksia ja tutkimussuuntia. Fei-Fei Li, Stanfordin yliopiston AI Labin johtaja, onkin verrannut tekoälyä ja sen teknologioita salaattikulhoon, joka pitää sisällään eri osa-alueita ja tieteenaloja. Tekoälyn koostuessa erilaisista tutkimussuunnista muuttuu myös tekoälyn ”sisä- ja ulkorajojen” määrittäminen haasteelliseksi. Mistä tekoäly alkaa ja mihin tekoäly päättyy? Rajanvedon ongelmista kertoo se, kuinka erilaisia julkaisumääriä eri määrittely- ja rajaustavoilla on mahdollista löytää. Maailmanlaajuisesti tekoälyä käsittelevien julkaisujen määrä vaihteleekin tuhansista lähes sataan tuhanteen julkaisuun vuodessa riippuen pelkästään aiheen rajauksesta.

Tekoälymenetelmien ja -tekniikoiden luonteesta johtuen, eron tekeminen soveltavan ja puhtaan teknologisen tutkimuksen välille on tekoälyn tapauksessa erittäin haastavaa. Esimerkiksi tekoälyn kehittämisessä keskeisiä algoritmeja tyypillisesti muokataan ja sovelletaan uusien tutkimuskohteiden ollessa kyseessä. Pelkästään nopeasti kehittyvän koneoppimisen alueella on nykyisin erotettavissa jo lähes 150 eri algoritmista menetelmää. Tässä tilanteessa puhe tekoälyn perustutkimuksesta ja soveltavasta tutkimuksesta toisistaan erillisinä tutkimuskonteksteina, joilla on omat tutkimusparadigmansa ja -menetelmänsä johtaa helposti ongelmallisiin tulkintoihin koko tekoälykäsitteen sisällöstä.

Bibliometetrisessä selvityksessä on lähdetty siitä, että yhtä ainoaa ja yleisesti hyväksyttävissä olevaa tapaa rajata ja ryhmitellä tekoälyn tutkimusta ei ole järkevää tavoitella. Julkaisuanalyysin menetelmät ja näkökulmat on esitetty tarkemmin liitteessä 2.

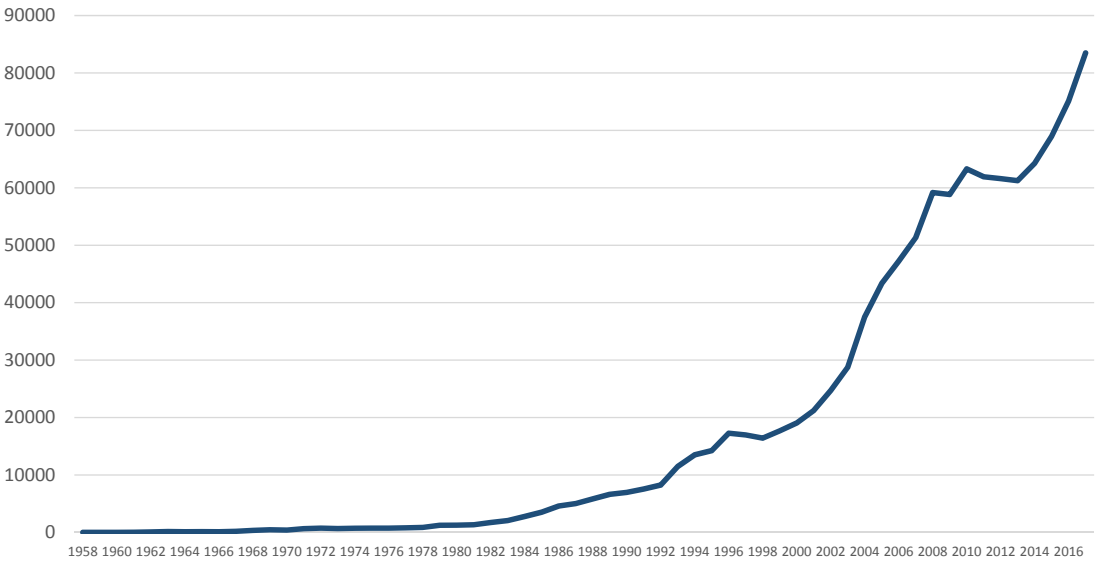
Julkaisukatsaus on kokonaisuudessaan luettavissa verkossa osoitteessa https://tietokayttoon.fi/documents/1927382/2158283/Teko%C3%A4lyn+huippuosaamisen+kartoitus/0f93a62e-f97a-464d-9946-ba3953914d04. Kyseinen laaja katsaus sisältää kaikkien kymmenen tutkimussuuntauksen julkaisuselvityksen lisäksi myös lyhyen katsauksen alan patentointiin. Tässä valtioneuvoston Tekoälyn kokonaiskuva -väliraportissa käsitellään tekoälyn huippuosaamista ja tutkimuksen vaikuttavuutta vain yhtenä kokonaisuutena. Yhteenvedossa esitetyt johtopäätökset pohjautuvat kuitenkin kaikkien kyseisten kymmenen tutkimussuunnan julkaisuanalyyseihin.

3.2 Julkaisuanalyysi

3.2.1 Julkaisumäärät ja aiheet

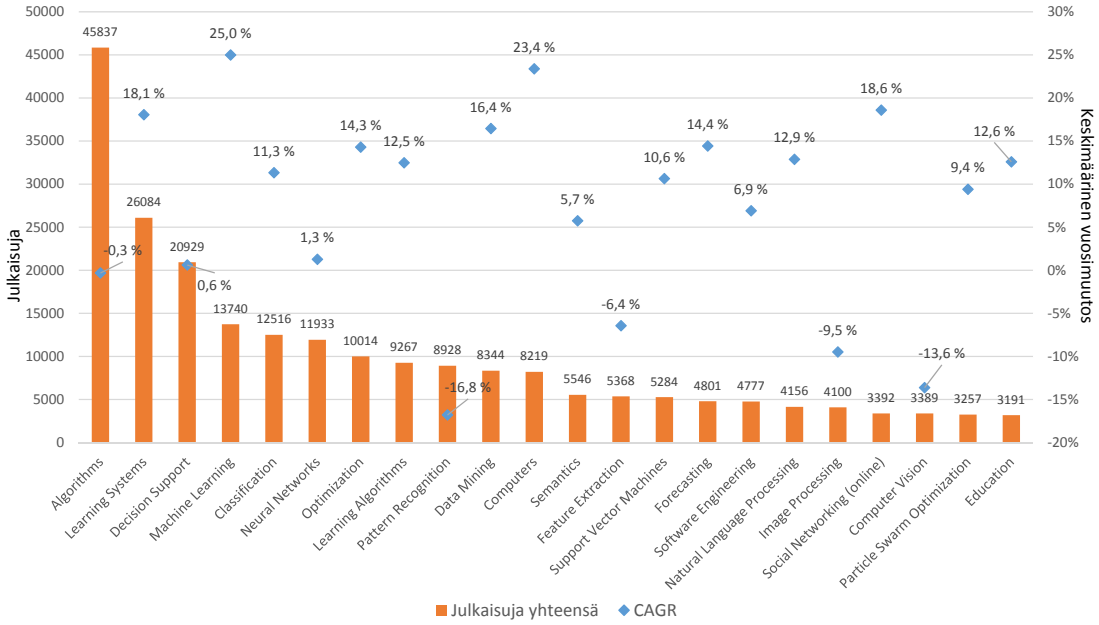
Tekoälyllä on tutkimusalana pitkät perinteet. Ensimmäiset tekoälyä ja sen mahdollisuuksia käsittelevät julkaisut ilmestyivät jo noin 60 vuotta sitten. Julkaisumäärien kasvu ei kuitenkaan ole ollut jatkuvaa, vaan väliin on mahtunut tasaisemman kehityksen ja jopa taantumisen kausia. Kuuluisin taantuma lienee ollut 1990-luvun ”tekoälyn talvi” (Crevier, 1993). Kokonaisuutena tekoälyn julkaisumäärät ovat kuitenkin kasvaneet 1960–2017 keskimäärin 19,2 % vuodessa. Mutta kasvunopeus on siis vaihdellut tänä aikana paljon. Nopean kasvun vaiheet ovat osuneet erityisesti tämän vuosituhannen puolelle.

Kuvio 10 Tekoälyä käsittelevien julkaisujen määrä 1958–2017



Lähde: Scopus (tilanne: 13.4.2018).

Kuvio 11 Tekoälyn keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu 2008–2017



Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 6 %.

Lähde: Scopus (tilanne: 5.3.2018).

Viime vuosien julkaisumääriä tarkasteltaessa nousevat aiheissa esiin erityisesti koneoppimiseen liittyvät teknologiat ja tutkimussuunnat ml. erilaiset neuroverkkoteknologioita ja -algoritmeja hyödyntävät ratkaisut. Vaikka esimerkiksi ns. syväoppiminen ja syvät neuroverkot ovat aivan muutamana viime vuonna nousseet erääksi nopeimmin kasvavaksi yksittäiseksi tutkimusalaksi, on ns. oppivien järjestelmien nousu tekoälytutkimuksen sisällä jatkunut jo pitkään. Toisaalta myös perinteiset aiheet ovat edelleen vahvasti esillä, kun tarkastellaan aiheiden julkaisumääriä useamman vuoden aikaperspektiivistä. Alla on poimintoja isoimmista teemoista ja nopeimmin kasvavista aiheista tekoälyjulkaisuissa.

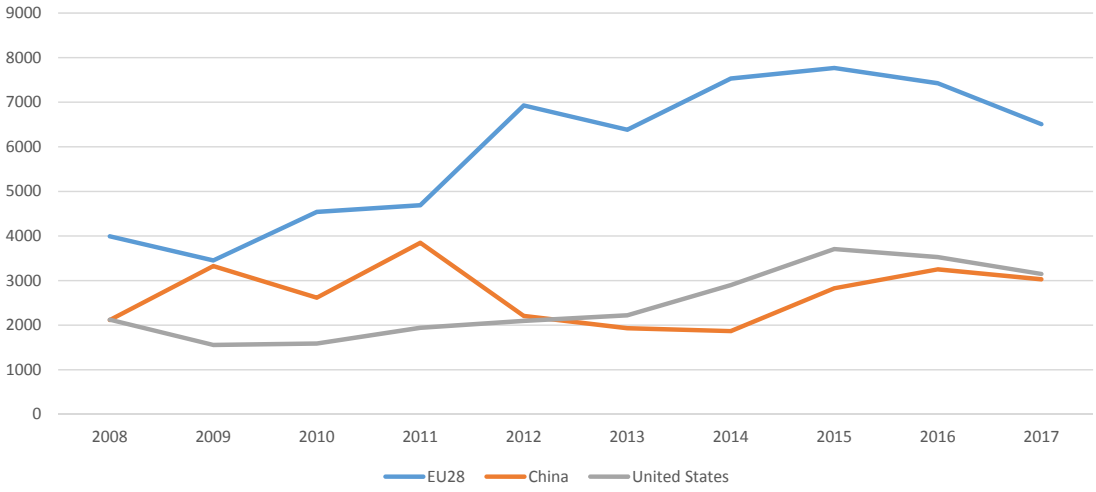
3.2.2 Maat

Kiina ja Yhdysvallat ovat tekoälyn suurvallat. Kiinan kasvavaa roolia kuvaa yleisemminkin se, että maa ohitti äskettäin Yhdysvallat maailman suurimpana tieteellisten julkaisujen tuottajana. Yhdessä kyseiset maat kattoivat v. 2016 n. 36 % kaikista tieteellis-teknillisistä julkaisuista.[7] Tekoälytutkimuksen piirissä Kiinan ja Yhdysvaltain osuus on vielä tätäkin merkittävämpi. Toisaalta osuus vaihtelee riippuen käytetystä tekoälykäsitteen rajauksesta. Esimerkiksi koneoppimisessa Kiinan ja Yhdysvaltojen osuus oli v. 2017 44 % kaikista julkaisuista. Kommunikaatiossa, havainnoinnissa ja toiminnassa osuus oli samana vuonna n. 39 %.

Tekoälytutkimuksen piirissä tilanne pysyi tarkastellulla ajanjaksolla suurimpien toimijoiden suhteen melko vakaana: kaikkien julkaisumäärät lisääntyivät. Suurinta vaihtelu oli Kiinan osalta. Vuosina 2014–2015 tapahtui selvä hyppäys julkaisumäärissä, mikä näyttäisi kuitenkin sittemmin tasaantuneen.

Taustalla näkyy ennen kaikkea datapohjaisten ja oppivien menetelmien lisääntynyt tutkimus.[8] Jo tällä hetkellä Kiinassa on eräs suurimmista tekoälytutkijoiden keskittymistä, ja on ennustettu, että Kiina ohittaa Yhdysvallat tekoälyn kehittämisessä vuonna 2025.[9] Tämä nä-

Kuvio 12 Julkaisumäärien kehitys: EU28, USA ja Kiina 2008–2017



Lähde: Scopus.

7 US National Science Foundation (NSF), Science & Engineering Indicators 2018. https://www.nsf.gov/statistics/2018/nsb20181/figures

8 Katso luku 2.3.6 Koneoppiminen.

9 Katso esimerkiksi The Economist, Marche 17th 2018, s. 11.

kyy myös julkaisutoiminnan kehityksessä. Selvästi nopeinta kasvu on kuitenkin ollut Intiassa sekä tarkasteltaessa koko ajanjaksoa 2003–2017 että vuosina 2008–2017. Toisaalta Intian lähtötaso on ollut huomattavasti Yhdysvaltoja ja Kiinaa matalampi.

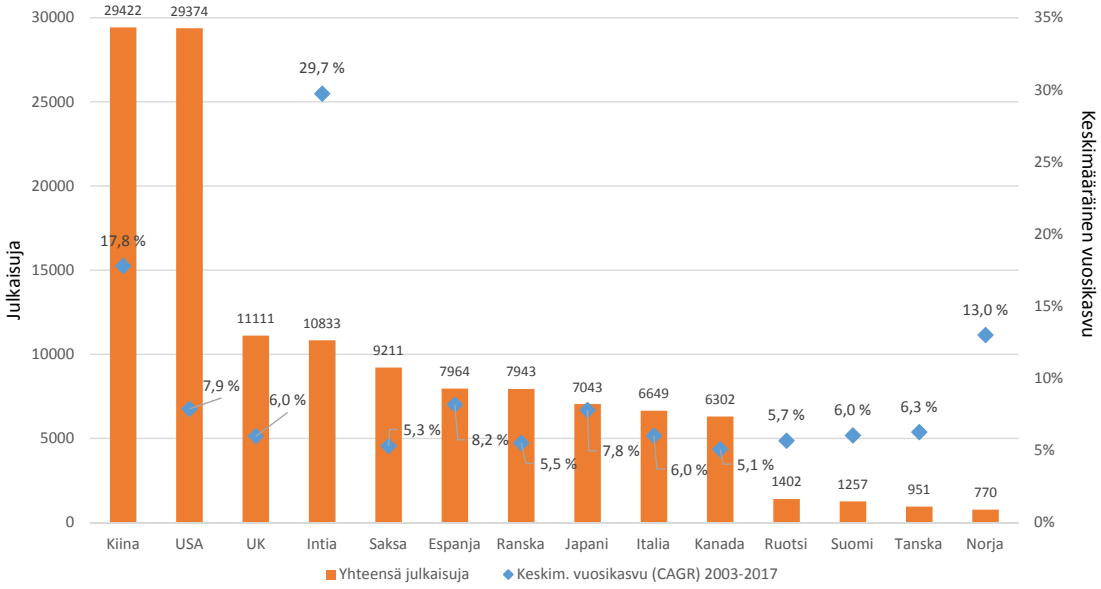
Mikäli tarkastellaan erikseen vain vuosia 2008–2017 ja otetaan mukaan julkaisuja väljemmällä hakurajauksella, niin Kiinan ja USA:n kasvunopeudet ovat ko. ajanjaksolla pienemmät. Muista maista Espanjan, Japanin ja Italian julkaisujen kasvu oli puolestaan nopeampaa.

Paitsi julkaisumäärissä ja tutkimuksen kasvunopeudessa on maiden välillä eroja myös tutkimuksen painotuksissa. Kuviossa 14 aiheiden prosenttiosuudet julkaisuista eri maissa vuosina 2015–2018.

3.2.3 Tutkimuksen vaikuttavuus: maat

Yhdysvallat on perinteisesti ollut joko kärjessä tai hyvin lähellä sitä tarkasteltaessa viittausmääriä millä tahansa tutkimusalalla. Toisaalta viittauksia keräävien kiinalaisten tekoälyjulkaisujen määrä on nykyisin jo 89 % amerikkalaisten julkaisujen määrästä[10]. Kiina on siten lähestymässä Yhdysvaltoja myös julkaisujen laadussa, mikäli julkaisun laadulla ymmärretään julkaisun keräämiä viittauksia muista julkaisuista. Julkaisun keskimäärin keräämien viitteiden määrässä Yhdysvallat on kuitenkin edelleen selvästi edellä Kiinaa.

Kuvio 13 Tekoäly: suurimmat maat ja Pohjoismaat 2003–2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus



Globaali keskiarvokasvu: 9,7 %.

Lähde: Scopus.

10 Chinese tech v American tech. The Economist, vol. 426 nr 9079, February 18th 2018, s. 61.

Kuvio 14 Tekoäly: suurimmat kymmenen maata ja Pohjoismaat 2015–2018, julkaisujen painotukset

Julkaisuja yhteensä:	10004	9213	5871	3403	2762	2388	2361	2301	2201	1847	1743	1250	468	372	295	248
Avainsana	USA	Kiina	Intia	UK	Saksa	Espanja	Ranska	Japani	Italia	Kanada	Australia	Etelä-Korea	Ruotsi	Suomi	Tanska	Norja
Algorithms	9,1 %	14,9 %	12,1 %	7,9 %	6,0 %	7,7 %	7,6 %	5,3 %	5,0 %	10,4 %	9,9 %	6,6 %	8,3 %	8,1 %	7,1 %	6,9 %
Big Data	5,0 %	4,7 %	4,7 %	3,7 %	3,1 %	3,2 %	2,5 %	3,3 %	4,5 %	3,8 %	4,8 %	4,3 %	3,6 %	4,6 %	2,0 %	5,2 %
Classification (of Information)	7,7 %	7,6 %	11,4 %	7,9 %	6,4 %	7,7 %	6,0 %	4,8 %	7,1 %	9,3 %	9,4 %	6,2 %	5,6 %	8,6 %	6,1 %	7,3 %
Computation Theory	3,3 %	4,5 %	6,3 %	3,7 %	2,3 %	2,8 %	3,4 %	2,5 %	4,3 %	2,9 %	3,0 %	1,3 %	2,6 %	1,3 %	2,7 %	4,4 %
Computer Science	4,1 %	4,3 %	1,4 %	7,7 %	13,1 %	4,5 %	12,1 %	10,0 %	7,7 %	5,6 %	4,6 %	2,5 %	10,3 %	11,0 %	16,6 %	11,7 %
Computers	5,9 %	5,5 %	1,9 %	10,7 %	18,6 %	6,0 %	17,6 %	12,9 %	10,7 %	8,0 %	6,4 %	2,9 %	14,5 %	15,1 %	26,1 %	15,7 %
Data Mining	8,5 %	7,0 %	17,1 %	8,1 %	7,7 %	8,0 %	6,8 %	6,6 %	7,5 %	8,8 %	10,7 %	6,9 %	7,3 %	8,6 %	1,7 %	4,8 %
Decision Making	3,9 %	2,7 %	3,5 %	4,9 %	3,9 %	5,8 %	5,3 %	2,3 %	5,9 %	4,3 %	5,9 %	3,0 %	4,9 %	4,8 %	2,0 %	5,6 %
Decision Support Systems	3,7 %	2,1 %	3,5 %	5,3 %	5,4 %	6,8 %	5,9 %	1,4 %	8,3 %	4,4 %	5,3 %	2,9 %	10,5 %	4,8 %	4,7 %	15,7 %
Decision Trees	3,6 %	2,4 %	4,7 %	4,1 %	2,6 %	3,8 %	2,9 %	3,1 %	3,2 %	4,5 %	3,8 %	3,5 %	5,1 %	2,7 %	1,0 %	3,6 %
Education	5,4 %	3,0 %	2,3 %	2,9 %	2,1 %	4,4 %	1,6 %	4,7 %	1,7 %	2,9 %	3,4 %	2,6 %	3,6 %	3,2 %	1,7 %	3,2 %
Feature Extraction	3,2 %	5,1 %	6,9 %	4,2 %	2,6 %	4,5 %	1,9 %	2,9 %	3,0 %	3,1 %	4,0 %	4,4 %	2,1 %	2,2 %	3,4 %	4,4 %
Forecasting	6,2 %	5,1 %	5,8 %	4,7 %	4,0 %	5,5 %	2,8 %	4,1 %	3,6 %	6,0 %	5,0 %	4,3 %	7,1 %	4,8 %	5,8 %	4,4 %
Learning Algorithms	14,6 %	10,4 %	12,6 %	10,0 %	9,5 %	8,1 %	8,0 %	8,2 %	8,4 %	13,0 %	10,4 %	10,2 %	9,2 %	7,5 %	8,5 %	8,1 %
Machine Learning	15,4 %	8,5 %	10,6 %	12,3 %	11,9 %	13,3 %	8,6 %	11,7 %	9,8 %	14,4 %	9,1 %	13,6 %	15,8 %	15,6 %	8,8 %	14,9 %
Neural Networks	5,0 %	5,8 %	6,8 %	7,2 %	6,8 %	11,6 %	5,3 %	6,3 %	6,0 %	6,2 %	5,6 %	6,5 %	3,6 %	7,8 %	3,7 %	3,6 %
Optimization	7,5 %	14,4 %	12,3 %	8,3 %	5,8 %	7,6 %	6,9 %	6,3 %	6,5 %	7,9 %	11,6 %	5,4 %	7,3 %	6,2 %	6,1 %	12,1 %
Pattern Recognition	2,5 %	3,0 %	4,8 %	3,6 %	2,6 %	3,3 %	2,5 %	2,5 %	3,0 %	3,2 %	3,2 %	2,7 %	2,6 %	6,2 %	2,0 %	
Semantics	4,0 %	5,2 %	2,6 %	4,4 %	5,4 %	3,9 %	5,8 %	3,8 %	7,3 %	3,5 %	4,5 %	2,3 %	5,8 %	5,1 %	4,4 %	2,8 %
Social Networking (online)	3,3 %	2,7 %	4,1 %	2,6 %	2,5 %	3,0 %	1,9 %	2,1 %	3,7 %	3,2 %	3,7 %	3,1 %	3,2 %	2,4 %	2,0 %	2,4 %
Support Vector Machines	3,4 %	4,6 %	6,4 %	3,0 %	2,4 %	2,7 %	1,8 %	3,7 %	2,9 %	4,0 %	3,0 %	3,4 %	1,7 %		2,0 %	2,8 %

Lähde: Scopus (tilanne: 27.3.2018).

Kuviossa 15 on listattuna tekoälytutkimuksen (rajaus: Elsevier SciValin research area -määrittelyä: ”artitificial intelligence”) merkittävimmät maat normalisoidun (suhteutetun / skaalatun) viittausindeksin (FWCI) perusteella. Raja-arvona tässä on FWCI:n arvo 1,08 (eli jos FWCI-arvo on 1,08, niin maan tekoälyä käsitteleviin julkaisuihin viitataan keskimäärin 8 % enemmän kuin koko maailman keskiarvon mukaan olisi odotettavissa). Alle tuhat julkaisua v. 2012–2017 tuottaneet maat on jätetty vertailusta pois.

Normalisoidulla viittausmäärällä mitaten tekoälytutkimuksen kolme kärkimaata ovat Singapore, Sveitsi ja Saudi-Arabia. Yksittäisiä huippujulkaisuja oli puolestaan suhteellisesti eniten Hongkongilla, Slovakialla ja Singaporella. Seuraavina tulivat Israel, Australia ja Kanada. Kansainvälistä yhteistyötä oli eniten Hongkongilla ja Saudi-Arabialla, molemmilla yli kaksinkertaisesti kansainväliseen keskiarvoon verrattuna. Yritysten ja yliopistojen välinen tutkimusyhteistyö oli kaikkein tiiveintä Israelissa, Irlannissa ja Yhdysvalloissa. Seuraavina tulivat Ruotsi, Sveitsi ja Saksa.

Kuvio 15 Tekoälytutkimuksen maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2012–2017

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehtirankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Singapore	2,26	3141	34,8	1,64	3,7	1,35
Switzerland	2,07	1993	22,9	1,64	6,3	1,36
Saudi Arabia	2,01	1603	26,7	2,01	1,2	1,67
Hong Kong	1,94	2370	41,2	1,99	3,2	1,45
Australia	1,73	6132	30,6	1,51	5,2	1,43
United States	1,69	30792	26,4	1,02	6,9	1,1
United Kingdom	1,63	11877	25,3	1,53	3,8	1,3
Austria	1,5	1522	23	1,47	4,1	1,21
Israel	1,5	1506	31,4	1,28	9,9	0,84
Italy	1,49	7484	18,8	1,13	2,6	1,77
Canada	1,49	5933	27,4	1,42	2,6	1,2
Pakistan	1,48	1275	11,4	1,26	0,6	1,71
Netherlands	1,46	2648	20,6	1,46	3,7	1,25
Finland	1,45	1291	19	1,41	3,6	1,45
Germany	1,44	9699	17,3	1,12	5,3	1,06
Belgium	1,4	1637	26,1	1,59	3,7	1,27
Spain	1,33	7335	28,3	1,16	2,3	1,46
Greece	1,25	1804	17	1,11	0,8	1,57
France	1,23	7672	19,4	1,39	4,6	1,04
Malaysia	1,19	2964	17,6	0,98	0,9	2,44
Egypt	1,16	1001	9,6	1,32	1,5	1,06
Poland	1,14	4042	17,6	0,77	0,7	1,05
Sweden	1,14	1739	13,2	1,5	6,7	1,26
China	1,11	49044	21,9	0,64	1,7	0,93

Viittausindeksin raja-arvo kuviossa: 1,08. Alle 1 000 julkaisua tuottaneet maat eivät ole mukana taulukossa.

Lähde: Scopus ja SciVal (tilanne: 4.5.2018).

3.2.4 Tutkimusorganisaatiot

Kiinan ja Yhdysvaltojen lisäksi Ranskan ja Ison-Britannian yliopistot ja tutkimuslaitokset ovat keskeisiä toimijoita tekoälytutkimuksessa. Ranskan kansallinen tutkimuskeskus CNRS, Centre national de la recherche scientifique, on nopeimmin tekoälyn tutkimustoimintaansa julkaisujen määrällä mitattuna lisännyt tutkimuslaitos. Kuten useilla muillakin tieteen aloilla Kiinan tiedeakatemia CAS on myös tekoälytutkimuksen alueella johtava toimija.

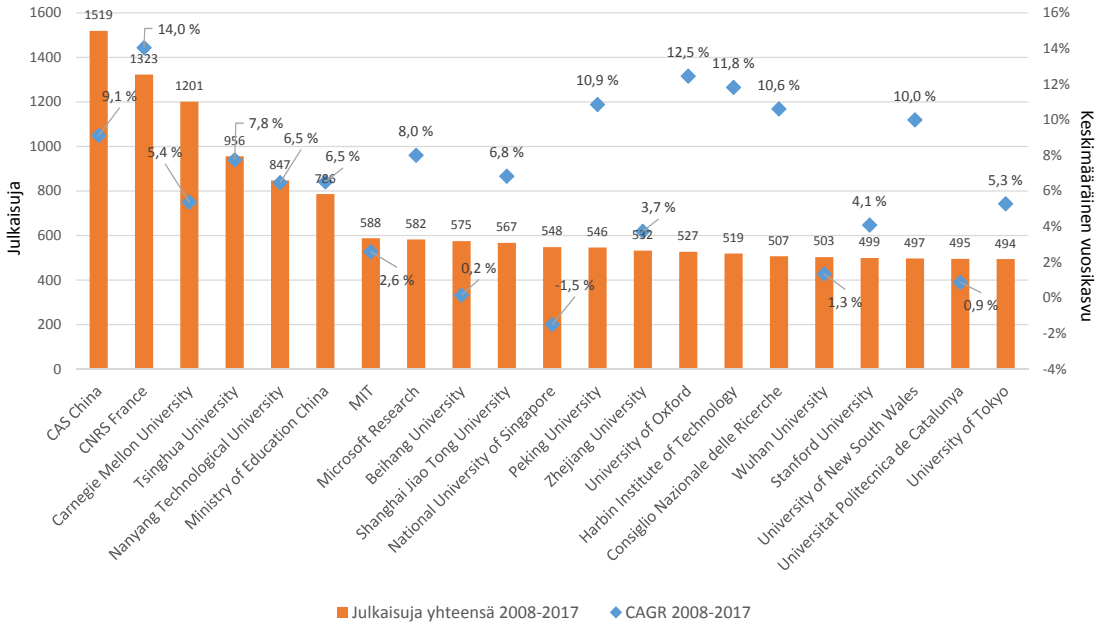
Suurimmat tutkimusorganisaatiot olivat tarkasteltavana ajanjaksona Kiinan tiedeakatemia ja Ranskan kansallinen tutkimuskeskus CNRS.

Euroopan suurimmat yliopistot tekoälyn julkaisumäärillä mitaten ovat viimeisen kymmenen vuoden aikana olleet Oxfordin yliopisto, Katalonian teknillinen yliopisto ja Leuvenin kuninkaallinen yliopisto. Tutkimuskeskuksista merkittävimmät ovat olleet CNRS:n lisäksi Italian CNR ja Ranskan INRIA. Suomessa suurimmat yliopistot ovat v. 2008–2017 odotetusti olleet Aalto yliopisto, Helsingin yliopisto ja Oulun yliopisto (Digibarometri, 2017).

Selkeästi suurimmat tutkimusjulkaisuja tuottavat yritykset v. 2008–2017 olivat amerikkalaiset Microsoft ja IBM. Julkaisumäärissä mitattuna nopeimmin panostustaan julkiseen tutkimukseen ovat lisänneet kuitenkin kiinalaiset Tencent, Baidu ja Alibaba Group sekä yhdysvaltalaiset Amazon ja Facebook. Julkaisumäärät ovat näillä kaikilla kuitenkin vielä toistaiseksi suhteellisen pieniä.

Koneoppimiseen, oppiviin algoritmeihin ja muihin oppiviin järjestelmiin liittyvän tutkimuksen osuus tekoälyn tutkimuksessa on nykyisin merkittävä yliopistoissa ja tutkimuslaitoksissa.

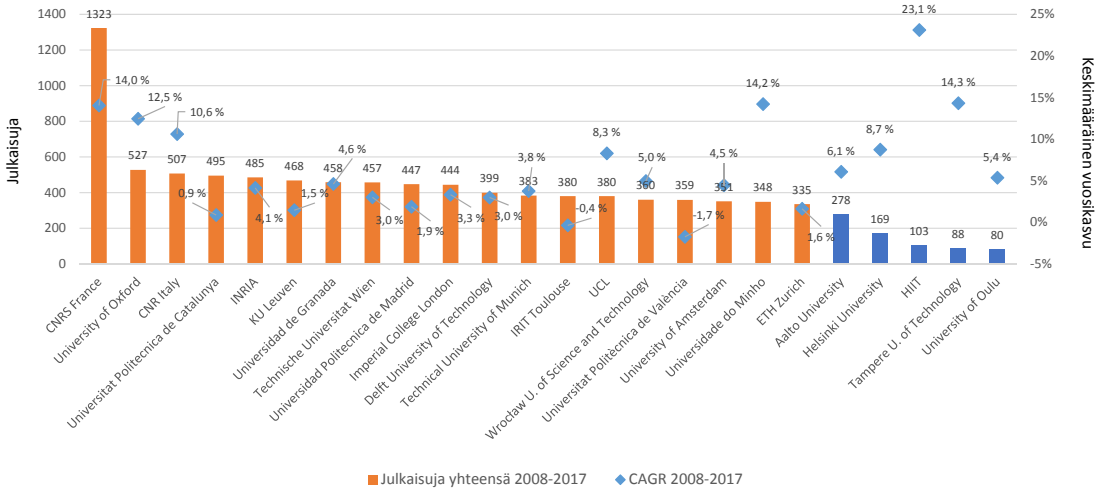
Kuvio 16 Tekoäly yleisesti tulkittuna: suurimmat tutkimusorganisaatiot 2008–2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus



Globaali kasvu 2008–2017: 5,7 %.

Lähde: Scopus (tilanne: 27.3.2018).

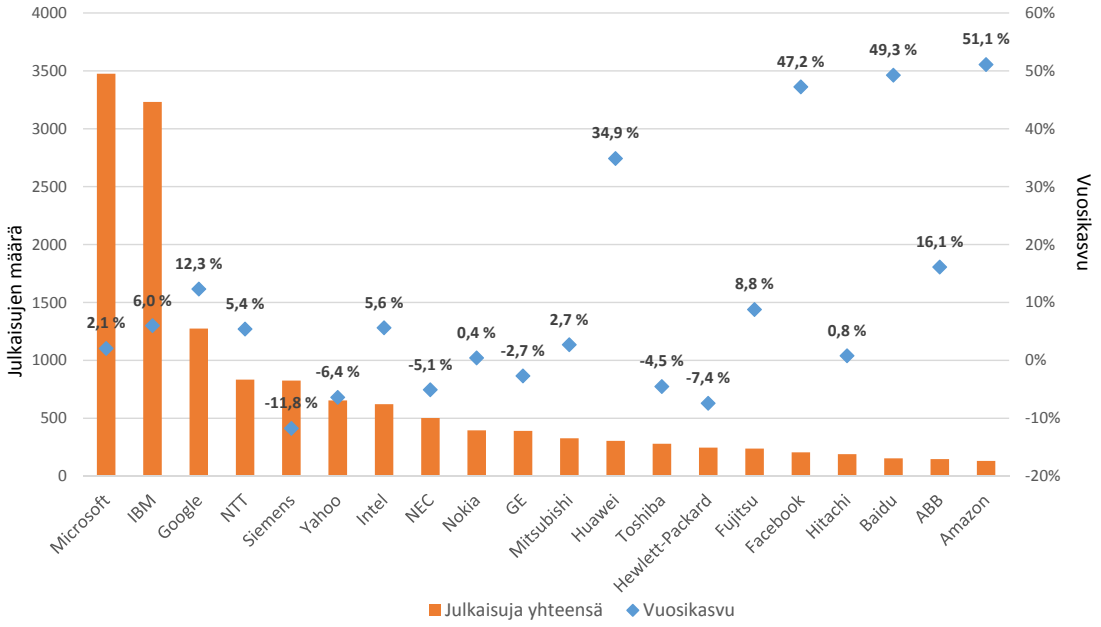
Kuvio 17 Tekoäly yleisesti tulkittuna: suurimmat eurooppalaiset ja suomalaiset tutkimusorganisaatiot 2008–2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus



Globaali kasvu 2008–2017: 5,7 %.

Lähde: Scopus.

Kuvio 18 Tekoäly yleisesti tulkittuna: suurimmat yritykset 2008–2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus



Lähde: Scopus.

Kuvio 19 kuvaa ennen kaikkea sitä, miten vahvasti oppivat järjestelmät nähdään osana tekoälyn kenttää. Suuri osuus kertoo lähinnä tekoälyn tutkimuksen perinteistä kyseisessä maassa: oppivat järjestelmät nähdään kyseisissä maissa todennäköisesti luontevammin osana tekoälyn tutkimuksen pitkää traditiota kuin siitä erillisenä tutkimussuuntana. Kuviossa on erotettu kymmenen johtavaa yliopistoa / tutkimuslaitosta, eniten julkaisevat yritykset ja eniten tekoälyjulkaisuja tuottaneet suomalaiset yliopistot v. 2015–2018.

Paitsi julkaisumäärien kasvunopeudessa on maiden välillä eroja myös tutkimuksen painotuksissa. Oppivat järjestelmät ja ohjelmistot ovat vahvasti edustettuna tutkimusaiheissa. Kuviossa 20 ovat aiheiden prosenttiosuudet julkaisuista johtavilla organisaatiolla vuosina 2015–2018.

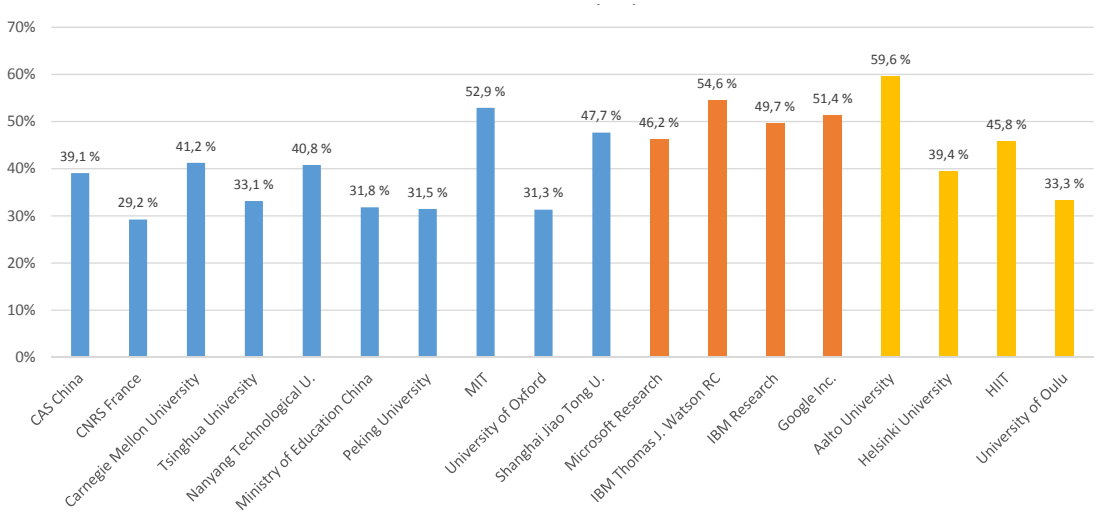
3.2.5 Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Yhdysvaltalaiset yliopistot ovat odotetusti vahvasti edustettuna, kun tarkastellaan tutkimuksen tieteellistä merkitystä. Eurooppalaisista toimijoista Zürichin ETH ja Oxfordin yliopisto ovat lähellä kärkeä käytettäessä mittarina julkaisujen keräämää keskimääräistä ja normalisoitua (odotettavissa olevaan viittausten keskiarvoon suhteutettua) viittausmäärää.

Yrityksen (Microsoft) asema tämän tekoälyä koskevan vertailun kärjessä on poikkeuksellista tutkimusjulkaisujen merkittävyyttä arvioitaessa. Osittain tämä tosin selittyy tiiviillä yhteistyöllä alan huippuyliopistojen kanssa. On silti hyvä todeta, että Microsoftin julkaisumäärät tekoälyn alalla ovat verrannolliset monien isojen yliopistojen julkaisumääriin.

Erilaisten tutkimuksen vaikuttavuutta kuvaavien mittareiden rinnalla on syytä aina tarkastella myös julkaisuvolyymejä. Vaikuttavuusmittarit sellaisenaan voivat antaa hyvinkin vinoutuneen kuvan toimijan (tai maan) tutkimuksen tasosta, mikäli tutkimussuunnan julkaisumäärä on pieni.

Kuvio 19 Oppivia järjestelmiä käsittelevien julkaisujen osuus kaikista tekoälyä käsittelevistä julkaisuista 2015–2018



Lähde: Scopus (tilanne: 27.3.2018).

Kuvio 20 Tekoäly yleisesti: suurimmat kansainväliset tutkimuslaitokset ja yritykset sekä ja suomalaiset yliopistot 2015–2018

Avainsana	699 CAS China	582 CNRS	565 Carnegie Mellon U.	483 Tsinghua University	353 Nanyang Technological U.	349 Ministry of Education China	286 Peking University	276 MIT	262 University of Oxford	256 Shanghai Jiao Tong U.	277 Microsoft Research	163 IBM Thomas J. Watson RC	147 IBM Research	142 Google Inc.	114 Aalto University	66 Helsinki University	48 HIIT	36 University of Oulu
Algorithms	9,6 %	6,7 %	6,4 %	9,5 %	10,8 %	20,9 %	9,8 %	9,8 %	7,3 %	9,8 %	7,2 %	9,8 %	11,6 %	9,2 %	9,6 %	6,1 %	8,3 %	11,1 %
Big Data	4,6 %	2,9 %	4,1 %	5,4 %	4,5 %	3,2 %	3,1 %	3,6 %	3,8 %	6,3 %	2,9 %	8,6 %	6,1 %	4,9 %	4,4 %	1,5 %		8,3 %
Classification (of Information)	8,3 %	5,8 %	4,1 %	7,0 %	10,2 %	9,5 %	6,3 %	5,4 %	5,3 %	9,0 %	5,8 %	6,1 %	4,8 %	1,4 %	10,5 %	1,5 %	8,3 %	13,9 %
Computation Theory	4,9 %	3,1 %	2,5 %	1,7 %	5,1 %	4,3 %		5,8 %	4,6 %	2,7 %	4,7 %			2,1 %				
Computer Science	5,3 %	16,7 %	3,0 %	6,4 %	3,1 %	5,2 %	11,5 %	5,4 %	7,6 %	3,5 %	6,5 %	3,7 %	2,0 %	2,1 %	6,1 %	13,6 %	10,4 %	5,6 %
Computer Vision	5,3 %	2,4 %	4,2 %	3,5 %	1,4 %	3,2 %	3,8 %	1,4 %	2,3 %	4,3 %	4,0 %			2,1 %			2,1 %	5,6 %
Computers	7,0 %	22,2 %	4,2 %	8,3 %	5,1 %	6,3 %	12,2 %	6,5 %	9,9 %	5,9 %	7,6 %	4,3 %	3,4 %	3,5 %	10,5 %	18,2 %	14,6 %	5,6 %
Data Mining	7,0 %	6,0 %	6,0 %	6,2 %	9,3 %	6,0 %	7,3 %	5,1 %	3,4 %	5,5 %	7,2 %	7,4 %	5,4 %	9,2 %	5,3 %	9,1 %	6,3 %	13,9 %
Decision Making	2,1 %	4,0 %	3,0 %	2,9 %	3,4 %	2,0 %	1,7 %	3,3 %	3,8 %	2,3 %	2,9 %	3,7 %	4,8 %	1,4 %	2,6 %			5,6 %
Feature Extraction	7,6 %	1,2 %	1,2 %	4,3 %	4,5 %	8,3 %	3,8 %	1,1 %	1,5 %	3,5 %		3,7 %	2,7 %	1,4 %	3,5 %			
Forecasting	6,9 %	1,4 %	4,6 %	5,4 %	4,0 %	2,6 %	4,9 %	8,0 %	1,9 %	8,6 %	7,2 %	9,2 %	13,6 %	2,8 %	7,9 %	3,0 %	4,2 %	8,3 %
Iterative Methods	4,0 %	2,4 %	4,2 %	2,9 %	2,8 %	2,9 %	4,5 %	5,4 %	1,1 %	6,6 %	2,9 %	3,1 %	2,7 %	3,5 %		3,0 %		
Learning Algorithms	13,6 %	5,7 %	13,1 %	9,7 %	17,3 %	11,7 %	9,8 %	13,8 %	6,9 %	11,3 %	13,4 %	18,4 %	11,6 %	8,5 %	8,8 %	9,1 %	10,4 %	8,3 %
Machine Learning	12,2 %	6,7 %	10,4 %	9,3 %	13,0 %	8,3 %	8,0 %	12,0 %	6,9 %	12,5 %	7,9 %	11,7 %	9,5 %	11,3 %	17,5 %	6,1 %	4,2 %	11,1 %
Matrix Algebra	2,9 %	2,9 %	2,8 %	2,5 %	2,3 %	2,3 %	3,1 %	4,7 %		4,3 %	4,0 %	3,7 %	4,1 %	2,1 %	4,4 %		4,2 %	
Natural Language Processing Sys.	4,0 %	1,4 %	2,1 %	2,9 %	5,4 %	2,3 %	7,0 %	2,2 %			4,7 %	4,3 %	3,4 %	4,9 %				
Neural Networks	11,6 %	2,9 %	1,2 %	5,6 %	7,1 %	4,9 %	4,9 %	4,0 %	2,7 %	5,5 %	4,3 %	3,7 %	4,1 %	4,2 %	7,0 %	4,5 %	4,2 %	5,6 %
Optimization	8,4 %	7,4 %	7,3 %	7,7 %	11,0 %	17,5 %	14,3 %	9,4 %	3,8 %	9,8 %	10,8 %	9,2 %	6,1 %	8,5 %	7,0 %	9,1 %	8,3 %	
Semantics	6,0 %	8,1 %	5,1 %	9,7 %	7,4 %	4,6 %	10,5 %	2,9 %	8,0 %	5,9 %	5,1 %	7,4 %	5,4 %	5,6 %		10,6 %	10,4 %	
Social Networking (online)	4,7 %	1,2 %	2,1 %	6,0 %	3,1 %	3,4 %	3,8 %	2,2 %	1,9 %	2,0 %	3,6 %	2,5 %			4,4 %		8,3 %	

Lähde: Scopus. Vuoden 2018 (tilanne: 27.3.2018).

Kuvio 21 Tekoäly: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2012–2017

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Microsoft USA	4,55	713	45	1,39	87,8	1,1
Stanford University	3,61	619	37,3	0,9	10,5	1,24
King Abdulaziz University	3,43	586	36,7	2,24	0,9	1,41
University of California at Berkeley	3,29	594	34,2	1,03	10,9	1,2
MIT	3,17	1180	28,8	1,1	9,1	1,41
ETH Zurich	2,85	631	26,1	1,63	7,6	1,33
University of Oxford	2,79	674	28	1,57	3,4	1,18
U. of Illinois at Urbana-Champaign	2,5	534	38,2	1,03	11,4	1,2
University of Michigan	2,49	492	-	0,73	9,6	1,31
Nanyang Technological University	2,48	1526	31,2	1,59	4,1	1,42
University of Granada	2,4	575	44,5	1,2	0,9	1,93
Harvard University	2,35	511	35	1,26	8	1,48
Carnegie Mellon University	2,33	1493	25,3	0,97	8,2	1,12
National University of Singapore	2,31	997	40,8	1,6	3,6	1,26
University of Toronto	2,28	516	28,3	1,44	6,8	1,29
Chinese University of Hong Kong	2,24	572	39,1	1,85	4,5	1,1
University of Technology Sydney	2,19	856	35,6	1,72	1,8	1,39
Institute of Automation CAS China	2,18	874	30,3	0,6	2,4	1,03
Hong Kong Polytechnic University	2,13	584	42,8	1,79	1,7	1,69
University of Southern California	2,1	702	24,7	0,94	6	1,09
Istituto Italiano di Tecnologia	2,05	506	-	1,32	2	1,55
University of Maryland	1,98	458	36,2	0,98	8,3	1,19
Imperial College London	1,96	631	24,9	1,58	2,7	1,34
EPFL Lausanne	1,96	454	-	1,69	5,1	1,24
Peking University	1,92	775	29,7	0,96	7,4	1,06
INRIA France	1,9	909	26,1	1,34	4,4	1,11
Georgia Institute of Technology	1,88	870	20,8	0,93	9,4	1,2
Technical University of Munich	1,86	788	13,7	0,94	8,8	1,29
University of Rome La Sapienza	1,82	548	21,4	1,06	2	1,77
University of Texas at Austin	1,82	451	-	0,77	8,6	1,02
Huazhong U. of Science and Tech.	1,77	965	30,1	0,82	0,6	1,19
Xidian University	1,71	857	35	0,57	1,8	1,02
IBM	1,68	510	-	1,36	71,8	1,07

Lähde: Scopus ja SciVal (tilanne: 4.5.2018).

3.3 Yhteenveto

Tässä selvityksessä on tarkasteltu tekoälyn julkaisukenttää eri näkökulmista.[11] Tutkimussuunnat poikkeavat huomattavasti toisistaan sekä julkaisumäärien että niissä tapahtuvien muutosten osalta. Pienimmän – etiikkaan ja regulaatioon liittyvän – tutkimusalan julkaisumäärä oli vuosina 2000–2017 vain murto-osa suurimman, koneoppimisen, julkaisumäärästä.

Sovellettaessa laajaa tulkintaa tekoälyn sisällöstä voidaan sitä käsittelevien julkaisujen määrän arvioida kasvaneen 2000-luvulla n. 9 %:n vuosivauhtia. Tämä on hyvin lähellä tutkimusjulkaisujen vuosikasvua koko tiedekentässä.[12] Tekoälyn osalta kasvu on kuitenkin vaihdellut havainnoinnin ja tilannetietoisuuden 7 %:n ja tekoälyn laskentaympäristöjen, alustojen ja palveluiden n. 15 %:n keskimääräisen vuosikasvun välillä.

11 https://tietokayttoon.fi/documents/1927382/2158283/Teko%C3%A4lyn+huippuosaamisen+kartoitus/0f93a62e-f97a-464d-9946-ba3953914d04

12 Ks. http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html ja https://blog.frontiersin.org/2017/03/06/160-million-papers-and-counting-the-worlds-information-deluge/

Erityisen nopeaa julkaisumäärien kasvu on ollut aivan viime vuosina, 2014–2017. Nopeaa kasvua selittää erityisesti koneoppimiseen ml. ns. syväoppimiseen ja keinotekoisiin hermoverkkoihin liittyvä – erityisesti soveltava – tutkimus. Joitain merkkejä on kuitenkin siitä, että nopeimman kasvun vaihe saattaa olla taittumassa.

Tekoälyn tutkimukselle on tyypillistä soveltavan tutkimuksen vahva asema. Kysymykseen tekoälyn perustutkimuksen ja soveltavan tutkimuksen erottamisen hyödyllisyydestä ei tässä selvityksessä otettu suoraan kantaa. Esimerkiksi nopeimmin kasvavat tutkimusjulkaisujen aiheet viime aikoina, kuten edellä mainitut CNN-, DNN- ja RNN-hermoverkot, liittyvät usein valtaosaltaan pikemminkin tekoälyn soveltamiseen (lääke- ja terveystiede, liikenne-logistiikka, kiertotalous, ilmastotutkimus, ruokaketjun hallinta jne.) kuin tekoälyn tutkimukseen sinänsä. Tämä ei kuitenkaan tarkoita, etteikö myös kyseinen soveltava tutkimus silti voisi viedä myös tekoälyn perustutkimusta eteenpäin.

Kuvio 22 Tutkimussuunnan julkaisuvolyymit, osuus ja julkaisumäärän vuosittainen kasvu vuosina 2000–2017

Tekoälyn tutkimussuunta	Julkaisuja yhteensä 2000-2017	Osuus	Vuosittainen kasvu 2000-2017
Koneoppiminen	815257	30,8 %	11,5 %
Havainnointi ja tilannetietoisuus	565819	21,4 %	7,1 %
Data-analyysi	388298	14,7 %	13,1 %
Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus	272157	10,3 %	10,8 %
Robotiikka ja koneautomaatio	190918	7,2 %	11,4 %
Luonnollinen kieli ja kognitio	172268	6,5 %	7,2 %
Vuorovaikutus ihmisen kanssa	159287	6,0 %	13,3 %
Järjestelmätaso ja systeemivaikutukset	51933	2,0 %	10,4 %
Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit	28710	1,1 %	14,8 %
Etiikka, moraali, regulaatio ja lainsäädäntö	1425	0,1 %	11,9 %

Lähde: Scopus.

Kun tarkastellaan julkaisutuotannon kehitystä maittain ja toimijoittain, niin mikään yllätys ei ole, että kiinalaisten yliopistojen ja tutkimuslaitosten julkaisumäärät ovat lisääntyneet tarkasteltavana ajanjaksona nopeasti. Kiina ei kuitenkaan ole ainut tekoälytutkimustaan nopeasti kasvattanut maa. Nopeimmin tekoälyn tutkimuspanostuksia lisänneitä maita julkaisumäärillä mitaten löytyy etenkin Etelä- ja Kaakkois-Aasiasta: Intia, Bangladesh, Indonesia, Thaimaa, Vietnam ja Pakistan. Muita nopeasti julkista tutkimustoimintaansa vuoden 2012 jälkeen lisänneitä ja samalla julkaisuvolyymiltaan merkittäviä maita olivat mm. Venäjä, Etelä-Afrikka, Saudi-Arabia ja Turkki. Euroopan unionin maista esimerkiksi Puola on nykyisin jo merkittävä toimija sekä julkaisuvolyymillä että julkaisumäärien kasvunopeudella mitaten.

Yleensä maiden tutkimuspanostukset painottuvat eri lailla eri tekoälyn tutkimussuunnissa. Esimerkiksi Intia on noussut USA:n ja Kiinan jälkeen kolmanneksi suurimmaksi tekoälyn julkaisumaaksi erityisesti koneoppimiseen, data-analyysiin ja havainnointiin liittyvän tutkimustyön myötä. Useissa tekoälyn eri tutkimussuunnissa on havaittavissa vastaavaa uusien maiden esiin nousua.

Tutkimusorganisaatioiden tasolla Kiinan Tiedeakatemia CAS ja Ranskan kansallinen tutkimuskeskus CNRS ovat tekoälytutkimuksessa omassa kokoluokassaan. CAS ja CNRS toimivat kuitenkin omissa maissaan eräänlaisina tutkimuksen sateenvarjo-organisaatioina ja osittain myös tutkimuksen rahoittajina, joten niiden – samoin kuin esimerkiksi Italian CNR:n – vertaaminen yksittäisiin yliopistoihin, ei ole mielekästä.

Eräitä Kiinan ja Yhdysvaltojen suuria yliopistoja ja tutkimuslaitoksia lukuun ottamatta tutkimusorganisaatiot yleensä erikoistuvat omalle tutkimusalalleen. Näin on myös tekoälyn ollessa kyseessä. Jos kuitenkin katsotaan tekoälytutkimusta kokonaisuudessaan, niin kiinalaisten Tsinghua- ja Beihang-yliopistojen ja Harbinin Teknologiainstituutin lisäksi kannattaa nostaa esiin USA:n Carnegie Mellon -yliopisto ja MIT sekä Nanyangin teknillinen yliopisto Singaporessa. Euroopassa Universite Paris Saclay, Ranskan INRIA-tutkimuskeskus, Technische Universität München ja Oxfordin yliopisto on syytä mainita erityisen merkittävinä toimijoina omissa maissaan. Pohjoismaissa Aalto-yliopisto ja Teknologian tutkimuskeskus VTT olivat vuosina 2013–2017 korkeimmalla olevat tutkimusorganisaatiot, kun kriteerinä käytettiin huippujulkaisujen osuutta kaikista vertaisarvioiduista tutkimusjulkaisuista (Haku: Scopus avainsanalla ”artificial intelligence” ja julkaisut rajattu tietotekniikkaan; julkaisuja vertailussa mukana 78 889).

Julkaisumäärien lisäksi on tärkeää pyrkiä arvioimaan myös tutkimustulosten tieteellisiä ansioita. Julkaisujen tieteellistä vaikuttavuutta arvioitaessa on olennaista tarkastella rinnakkain eri mittareita. Käytettäessä mittarina suhteutettuja viittausmääriä olivat maista kärjessä useilla tutkimusaloilla mm. Singapore, Sveitsi ja Hongkong. Uusiin nouseviin maihin kuului Saudi-Arabia, jolla on julkaisumäärien lisäksi nykyisin näyttöä myös tutkimuksen tasosta. Suurista julkaisujen tuottajamaista Yhdysvallat, Iso-Britannia, Australia ja Kanada ovat myös julkaisujen vaikuttavuuden näkökulmasta (edelleen) merkittäviä toimijoita. Euroopassa Espanja on nykyään julkaisumääriltään Ranskan ja Italian kokoluokassa ja ohittanut ne esimerkiksi huippujulkaisujen osuudella mitaten.

Myöskään Kiina ei enää ole vahvasti esillä pelkästään julkaisumäärien takia, vaan yhä enemmän myös julkaisujen tieteellisen tason ansiosta. Tosin laatu näkyy vielä toistaiseksi ehkä enemmän yksittäisinä huippuina kuin tason kokonaisvaltaisena nousuna. Kiinan voimakas panostaminen tekoälytutkimukseen ja samalla maan lisääntyvä tutkimusyhteistyö selittänevät osaltaan myös Hongkongin ja Singaporen vahvan aseman tekoälyn huippututkimuksessa.

Eräs Suomen vahvuus tutkimuksen vaikuttavuuden näkökulmasta on ollut kansainvälisen yhteistyön suhteellisen suuri rooli tutkimusjulkaisujen tuottamisessa. Kansainvälisen tutkimusyhteistyön osuus oli merkittävimmistä tekoälyjulkaisujen tuottajamaista suurin Saudi-Arabiassa ja Hongkongissa, joissa molemmissa kansainvälisten yhteisjulkaisujen osuus oli kaksinkertainen verrattuna yhteisjulkaisujen kansainväliseen keskitasoon (Saudi-Arabiassa 75 % ja Hongkongissa 73 %). Maiden erona oli yhteistyön jakautuminen kumppanimaiden välillä: Hongkong teki yhteistyötä ennen kaikkea Kiinan kanssa, Saudi-Arabian yhteisjulkaisut jakaantuivat puolestaan tasaisemmin useiden maiden kesken.

Kansainvälisen tutkimusyhteistyön osuutta v. 2013–2017 tarkasteltaessa erottuivat muista maista erityisesti Singapore ja Sveitsi. Käytännössä kyse on siten hyvin pitkälle samoista maista, joiden julkaisut keräävät paljon viittauksia muista julkaisuista. Kansainvälisen tutkimusyhteistyön määrän ja erityisesti juuri viittausmäärien onkin havaittu korreloivan toisiaan.

EU-maista kansainvälistä julkaisuyhteistyötä oli tarkastelujaksona suhteellisesti eniten Tanskassa, Belgiassa, Irlannissa ja Isossa-Britanniassa. Myös Suomi ja muut Pohjoismaat olivat tässä vertailussa hyvin esillä, kansainvälisten yhteisjulkaisujen osuuden ollessa n. 50 % sekä Ruotsissa, Norjassa että Suomessa.

Myös yritys-yliopisto (tutkimuslaitos) -yhteistyö on Suomessa keskimääräistä korkeammalla tasolla, vaikka ei ylläkään aivan huippumaiden tasolle. Kyseisiä maita olivat tarkasteluajanjaksona Israel, Sveitsi ja Yhdysvallat. Euroopan unionin maista erityisesti Ruotsilla ja Irlannilla on vahvaa näyttöä yritys-yliopisto -yhteistyöstä eri tekoälyn tutkimussuunnissa. Suurista EU-maista erityisesti Saksassa on yritysten ja yliopistojen tutkimusyhteistyöllä vakiintunut rooli. On kuitenkin hyvä todeta, että yritysten ja tutkimusorganisaatioiden yhteisjulkaisujen osuus ei edes huippumaissa ylitä 10 %:tä kaikista julkaisuista.

Tutkimusorganisaatioiden vertaaminen keskenään tai listaaminen tutkimuksen vaikuttavuuden perusteella ei ollut tämän selvityksen ensisijaisena tavoitteena. Tekoälyn tutkimuksen tekee kuitenkin poikkeukselliseksi muutamien suurten yritysten, erityisesti Microsoftin, IBM:n ja osittain myös Googlen vahva panostus patentoinnin ohella myös julkiseen tutkimukseen. Esimerkiksi Microsoftin julkaisumäärät olivat v. 2008–2017 samaa luokkaa keskisuuren EU-maan kanssa, joten julkaisuja ja niihin liittyvää julkista tutkimusta ja tutkimusyhteistyötä ei voi pitää pelkkänä yrityksen t&k-toiminnan sivuvirtana, ainakaan suurimpien yritysten osalta. Microsoft sijoittui korkealle myös tarkasteltaessa julkaisujen tieteellistä vaikuttavuutta koko tekoälyn julkaisukentässä. Vastaavasti Google oli kärjessä koneoppimista käsittelevien julkaisujen viittausmäärissä. Jos lääketieteellistä tutkimusta ei tässä yhteydessä huomioida, niin yritysten sijoittumista korkealle myös julkaisujen vaikuttavuusvertailuissa voi pitää tieteellisen tutkimuksen piirissä poikkeuksellisena. Yritysten tiivis yhteistyö ja kumppanuus huippuyliopistojen kanssa selittänee tästä vain osan.

4 TEKOÄLYTEKNOLOGIOIDEN OMINAISUUDET JA KEHITTYMINEN

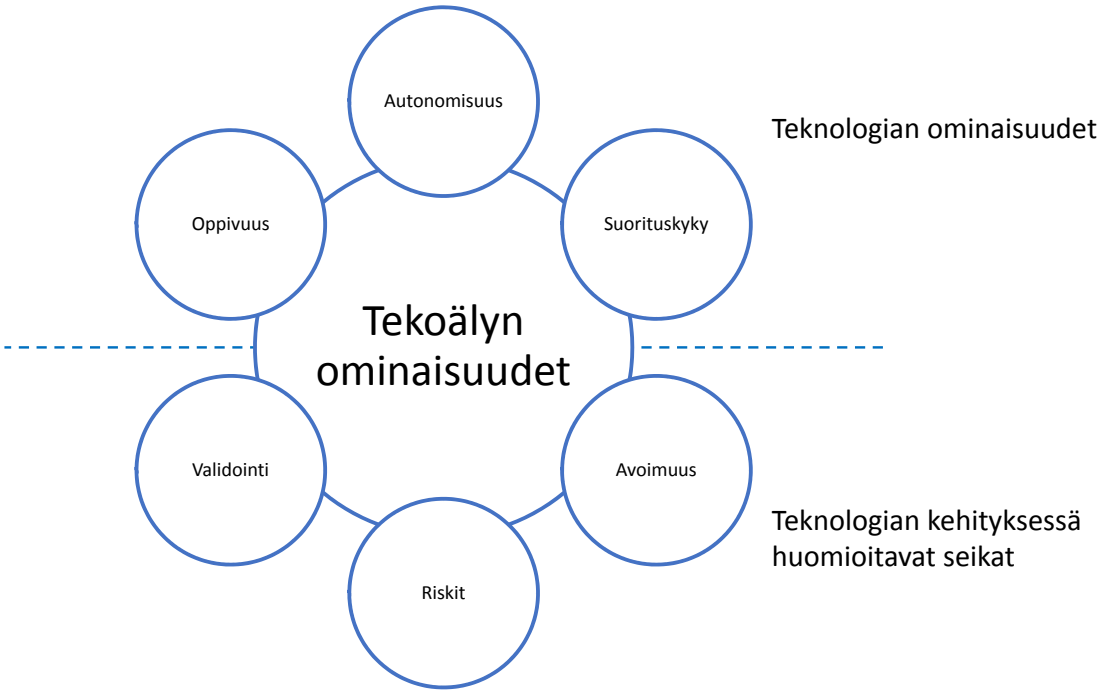
Raportissa on jo kuvattu tekoälyn yleisiä ominaispiirteitä, niihin liittyviä osaamisalueita sekä tekoälyn tutkimuksen kannalta olennaisia tietoaineistoja. Tässä luvussa paneudumme tarkemmin itse tekoälyteknologioihin, niiden ominaisuuksiin ja odotettavissa oleviin kehityssuuntiin. Aluksi erittelemme ominaisuuksia, joiden pohjalta teknologioita on tarkoituksenmukaista tarkastella. Tämän jälkeen arvioimme keskeisimpiä tekoälyteknologioita näiden ominaisuuksien näkökulmasta. Tarkastelussa huomioidaan sekä ohjelmistoteknologiat että tekoälyn hyödyntämisen kannalta oleelliset laitteistotekniset kehityspolut.

4.1 Teknologioiden ominaisuudet

Määrittelimme raportin alussa tekoälyn välineeksi, jonka avulla koneet, laitteet, ohjelmat, järjestelmät ja palvelut voivat toimia tehtävän ja tilanteen mukaisesti järkevällä tavalla. Järkevä toiminnan taso edellyttää tekoälyltä tiettyjä ominaisuuksia, sen on osattava tunnistaa erilaisia tilanteita ja ympäristöjä ja toimittava muuttuvien tilanteiden mukaan. Nämä ominaisuudet vaativat tekoälyltä autonomisuutta, oppivuutta ja suorituskykyä – tekoälyn on itse tunnistettava erilaisia tilanteita, osattava toimia ilman ennalta ohjelmointia ja suoriuduttava tehtävistä järkevällä tavalla. Kuvio 23 esittää tekoälyn ominaisuudet ja kehityksessä huomioon otettavat asiat.

Teknologioiden ominaisuudet liittyvät toisaalta teknologioiden kyvykkyyteen suoriutua tehokkaasti annetusta tehtävästä, toisaalta taas syvällisemmin oppivuuteen ja kykyyn toimia

Kuvio 23 Tekoälyn ominaisuudet, joihin sisältyy sekä valmiin teknologian ominaisuuksia että teknologioiden kehityksessä huomioitavia seikkoja



autonomisesti. Oppivuuden ja autonomisuuden mittaaminen on erityisen vaikeaa teknologian jatkuvasti kehittyessä. Samoin kuin tekoälylle, myöskään ominaisuuksille ei ole olemassa selkeää määrittelyä, mutta tehtävänannon mukaisesti tärkeimmiksi ominaisuuksiksi tässä raportissa luetaan seuraavat:

- **oppivuus** - pystyykö tekoäly oppimaan itse ilman ennalta ohjelmointia
- **suorituskyky** - millaisista tehtävistä tekoäly pystyy suoriutumaan esimerkiksi verrattuna ihmisen suorituskykyyn
- **autonomisuuden taso** - millä tasolla tekoäly pystyy itsenäiseen päättelyyn
- **avoimuus** - tekoälyn päätöksenteon läpinäkyvyys ja ennakoitavuus ja tekoälyn kehittämiseen käytetyn lähdekoodin ja menetelmien avoimuus.
- **riskit** - tekoälyn riskien arviointi liittyen toiminnan turvallisuuteen, yksityisyyteen ja tietoturvaan, hallittavuuteen ja läpinäkyvyyteen
- **validointi** - miten arvioidaan tekoälyn kehityksen onnistumista ja millä mittareilla

Pyrimme purkamaan tekoälyn ominaisuudet näiden attribuuttien mukaisesti helpommin ymmärrettäviksi. Samoin pyrimme ennakoimaan ominaisuuksien kehittymisen aikajänteen, mikäli sellaista tietoa on saatavilla.

Tekoälyn viimeaikainen nopea kehittyminen on tuonut mukanaan myös suuria toiveita ja pelottavia uhkakuvia tekoälyn ominaisuuksien suhteen. Tekoäly voi ominaisuuksiensa puolesta suoriutua hyvin tietyillä kapeilla sektoreilla, mutta laajasta ymmärryksestä (general AI) olemme vielä kaukana. Samalla on hyvä kuitenkin tiedostaa tekoälyn ominaisuuksien kehittyminen, joka voi kapeilla sovellusalueille on hyvinkin nopeaa.

4.1.1 Oppivuus

Oppivuudella tarkoitetaan tekoälyjärjestelmän kykyä toimia tehtävän tai tilanteen mukaisesti järkevällä tavalla muuttuvassa ja kompleksisessa ympäristössä. Käytännössä tämä vaatii tekoälyltä kyvykkyyttä oppia, koska kaikkia tilanteita ei voida etukäteen ohjelmoida, ellei sovellusalue ole tiukasti rajattu. Aiemmin tekoäly perustui usein sääntöpohjaisiin, ennalta ohjelmoituihin sääntöihin, koneoppimisen läpimurron myötä tekoäly pystyy oppimaan opetusaineistosta ja osaltaan tekemään päätelmiä myös opetusaineistoon kuulumattomista tapauksista.

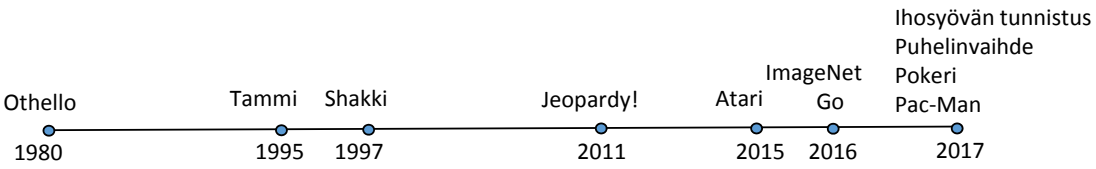
Nykyisissä tekoälyjärjestelmissä oppivuuden taso riippuu käytetystä menetelmästä. Ohjatussa oppimisessa ohjataan sekä opettamista että haluttua tavoitetta, vahvistetussa oppimisessa oppivuus perustuu vuorovaikutukseen ympäristö kanssa, syväoppimisessa oppiminen perustuu valtavaan data-aineistoon ja laskentatehokkuuteen

4.1.2 Suorituskyky

Tekoälyä verrataan usein suorituskyvyltään ihmiseen ja mietitään milloin tekoäly ylittää ihmisen suorituskyvyn. Tietyn osa-alueen tehtävissä tekoäly suoriutuu helposti ihmistä paremmin, kuten esimerkiksi hahmontunnistuksesta. Laajempaa päättelyä tai yhteyksien ymmärtämistä vaativat tehtävät ovat tekoälymenetelmille edelleen hyvin vaikeita.

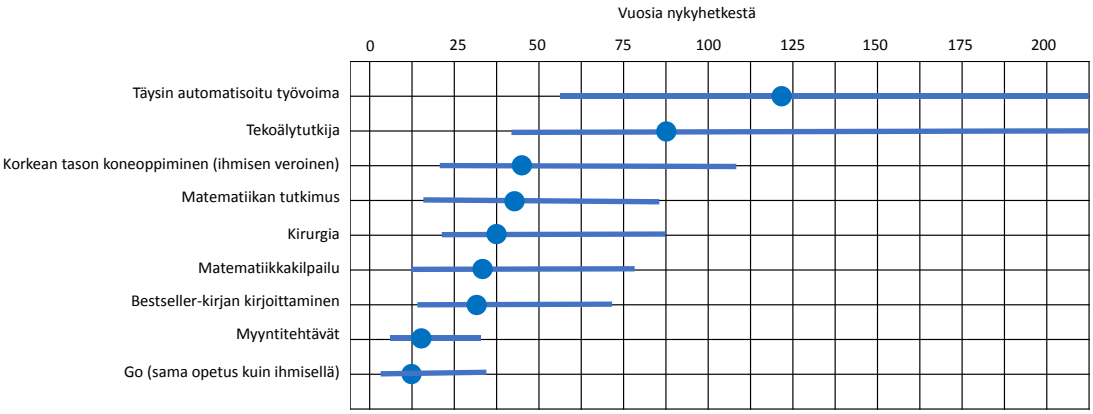
Ihminen voittaa yleisissä kysymyksissä tekoälyn vielä pitkään, mutta tietyillä kapeilla aihealueilla tekoälyn suorituskyky on jo saavuttanut tai ylittänyt ihmisen suorituskyvyn. Usein tekoälyn suorituskykyä verrataan ihmiseen mittaamalla tekoälyn kykyä pelata pelejä. Pelit ovat helposti hallittava, rajattu ympäristö testien suorittamiseen. Kuvio 24 (AI100, 2016) esittää merkkipaalut, jolloin tekoäly on voittanut ihmisen tietyssä pelissä.

Kuvio 24 Merkkipaalut vuosista, jolloin kone on voittanut ihmisen



Tekoälyn suorituskyky voidaan jakaa kahteen tasoon – teknologiseen suorituskykyyn ja yleiseen suorituskykyyn. Teknologisella suorituskyvyllä voidaan mitata esimerkiksi, milloin tekoäly tunnistaa kuvan, vastaa kysymyksiin tai tekee kielenkäännöksiä paremmin kuin ihminen. Laajemmassa merkityksessä mietitään usein, milloin tekoäly korvaa ihmisen työtehtävissä ja millaiset työtehtävät tekoäly pystyy korvaamaan. Termillä ”High-level machine intelligence” (HLMI) mitataan koneen kykyä suoriutua itsenäisesti jokaisesta työtehtävästä paremmin ja halvemmalla kuin ihminen. Kuvio 25 esittää tekoälyntutkijoille tehdyn kyselyn perusteella arvion aikajaksosta, jolloin tekoäly pystyy korvaamaan työntekijöitä eri tehtävissä. (Grace ym., 2018)

Kuvio 25 Tekoälytutkijoiden arvio koneen suorituskyvyn kehityksestä ihmiseen verrattuna



Arviot tämän tason saavuttamisesta vaihtelevat suuresti jopa tekoälytutkijoiden keskuudessa. Kapealla osaamisalueella tekoäly korvaa ihmisen lyhyellä aikavälillä, mutta osaamisissa, joissa tarvitaan laajempaa tietämystä ja osaamista, arvioiminen on mahdotonta.

4.1.3 Autonomisuuden taso

Autonomisuudella tarkoitetaan järjestelmän kykyä suorittaa tehtäviä ja saavuttaa tavoitteita kompleksisessa muuttuvassa ympäristössä ihmisen puolesta tai yhteistyössä ihmisen kanssa. Autonomisuus siis tarkoittaa järjestelmän kykyä itsenäiseen toimintaan. Nykyiset järjestelmät pystyvät ongelman ratkaisuun niille etukäteen opetetun aineiston perusteella, eli ne eivät vielä tässä vaiheessa ole autonomisia.

Tekoälyratkaisujen autonomisuuden tasot ulottuvat jäykähköstä ohjelmoidusta automaatiosta järjestelmän itsenäiseen päätöksentekoon seuraavasti:

– automaattinen älykkyys, manuaalisten ja rutiininomaisten töiden automatisointi (vähäinen autonomisuus)
– avustava älykkyys, järjestelmällä avustava rooli tehtävien suorittamisessa paremmin ja nopeammin (alhainen autonomisuus)
– lisätty älykkyys, järjestelmä auttaa tehtävien suorittamisessa esimerkiksi päätöksen teon tukena
– autonominen älykkyys, järjestelmä tekee automaattisesti päätöksiä ilman ihmisen puuttumista (korkea autonomisuus)

4.1.4 Avoimuus

Tekoälyteknologiaan liittyen avoimuus voi kattaa useita asioita: avointa lähdekoodia, tutkimustiedon avoimuutta, datan saatavuutta, turvallisuusteknologiaan liittyvää avoimuutta sekä yritysten tekoälyyn liittyviin kyvykkyyksiin ja tavoitteisiin liittyvää avoimuutta. Koska tekoälyllä on tulevaisuudessa merkittävä vaikutus jokapäiväiseen elämäämme ja koko yhteiskunnan toimintaan, avoimuudella on erittäin tärkeä rooli tekoälyn kehityksessä.

Suurimmat tekoälyn kehittäjätahot (Google, Facebook, Microsoft ja Baidu) ja akateeminen yhteisö jakavat tietoa konferensseissa ja verkossa, täydelliseen avoimuuteen on kuitenkin vielä matkaa. Esimerkiksi lähdekoodien laajempi avaaminen auttaisi ymmärtämään tekoälyn päätöksenteon perusteita ja lisäämään läpinäkyvyyttä. Avoimuuden edistäminen on tärkeää, ettei tekoälyteknologian osaaminen keskity vain harvojen käsiin, ja osaamista käytetään yhteiskuntaa edistäviin asioihin. (Bostrom, 2017, www.openAI.com)

4.1.5 Riskit

Viime aikoina tekoälykeskusteluun on noussut yhä enemmän keskustelua tekoälyn riskeistä, joskaan keskustelu ei liity itse tekoälyteknologiaan vaan enemmän siihen, miten ja missä tarkoituksessa tekoälyä käytetään. Tekoälyteknologialla on merkittävä rooli jo nykyisissä sovelluksissa, puheentunnistuksessa, hakukoneissa ja automaattisessa kääntämisessä, tekoälyn rooli kasvaa koko ajan tulevissa sovelluksissa kuten autonomisessa ajamisessa, työtehtävien avustamisessa ja etäoperoitavissa tehtävissä. Tietyt sovellusalueet korostuvat riskikeskustelussa, militäärisovellukset ja yksityisyyteen liittyvä sovellukset (esim. automaattinen kasvojentunnustus turvallisuussovelluksissa), joiden väärinkäytön mahdollisuudet pelottavat.

Itse tekoäly toimii sille määritellyllä tavalla, joten tekoälyn kehittäjillä on suuri vastuu. IEEE on julkaissut raportin tekoälyn ja autonomisten järjestelmien eettisen suunnittelu periaatteista (Ethically aligned design), joiden tavoitteena on tekoälyn teknologia, joka noudattaa ihmiskunnan moraalisia ja eettisiä arvoja (IEEE, 2016). Myös joukko merkittäviä tutkijoita on julkaissut ns. Asilomar-periaatteet (Asilomar AI principles) tekoälytutkijoiden johtamana (Asilomar, 2018).

Väärinkäytettynä tekoäly voi uhata digitaalista turvallisuutta (hakkerointi, datan väärinkäyttö), fyysistä turvallisuutta (autonomisten koneiden väärinkäyttö) ja poliittista turvallisuutta (yksityisten tietojen väärinkäyttö, poliittiset kampanjoinnit). Nämä laajamittaiset mahdolliset kielteiset vaikutukset on huomioitava tekoälyä suunniteltaessa, joka on aina osa laajempaa digitaalista infrastruktuuria (Malicious use of AI, 2018).

4.1.6 Validointi

Validoinnilla tarkoitetaan menettelyä, jossa tarkastellaan kehitettävän teknologian vaatimustenmukaisuutta. Tekoälyteknologiat ja niiden hyödyntäminen reaalimaailman sovelluksissa tuovat mukanaan uusia haasteita järjestelmien validoinnin eri vaiheisiin. Tekoälyä ja etenkin koneoppimista hyödyntävien järjestelmien validointiin liittyy joukko uudenlaisia kysymyksiä:

- Miten varmistetaan, että järjestelmän opettamisessa käytettävä data on laadukasta ja tarkoituksenmukaista sekä riittävän kattavaa?
- Miten varmistetaan, että järjestelmä oppii halutulla tavalla, eikä esimerkiksi vääränlaisia painotuksia pääse syntymään?
- Miten varaudutaan mahdollisiin muutoksiin järjestelmän käyttöympäristössä?
- Millaisilla menetelmillä yllä olevat tarkastelut tehdään, kun kyseessä on monimutkainen ja/tai dataintensiivinen järjestelmä?

Ohjelmistotekniikassa validoinnin suorittamiseen on olemassa standardoituja menettelytapoja, joiden avulla järjestelmää voidaan systemaattisesti arvioida (esim. IEEE Std 1012, 2016). Vastaavia menetelmäkuvauksia on olemassa myös mm. automaatiojärjestelmien validointiin, mutta erityisesti tekoälyteknologioihin keskittyvää ohjeistusta ei ole tällä hetkellä saatavilla. Teknologioiden kehittäjillä voidaan siis katsoa olevan suuri vastuu validoinnin kehittämisestä.

4.2 Keskeisten tekoälyteknologioiden ominaisuuksien arviointi

Kuten raportin aiemmista luvuista käy ilmi, on tekoälyn kenttä varsin laaja ja laveasti määritelty. Niinpä alan kirjallisuudessa on esitetty runsaasti erilaisia jaottelutapoja ja hierarkioita myös tekoälyteknologioiden jäsentämiseen (ks. myös liite 1). Jäsennys voi olla esimerkiksi keskeisiin tekoälyn matemaattisiin menetelmiin pohjautuva, vaikkapa algoritmityyppeihin perustuen. Toisaalta usein on hyödyllistä lähestyä teknologioita sovellusalueiden näkökulmasta, jolloin eri teknologioiden ominaisuuksien vertailu on hedelmällisempää.

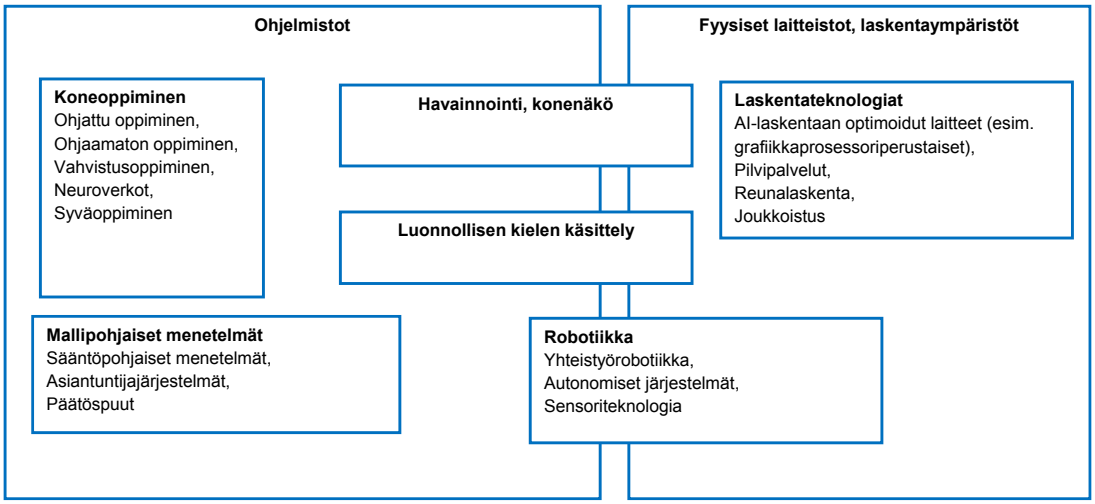
Tässä raportissa on hyödynnetty Stanfordin yliopiston ja MIT:n listauksia merkittävimmistä tutkimusalueista, jolloin voidaan keskittyä tällä hetkellä oleellisimpina pidettyihin teknologioihin. Tekoälyn osa-alueiden näkökulmasta käsittelemme teknologioiden ominaisuuksia seuraavasti:

1. **Data-analyysi** on tässä yhteydessä käsitetty menetelmäjoukkona, joka tukee varsinaisia tekoälyteknologioita, joten ominaisuuksia ei ole erikseen eritelty.
2. **Havainnointi ja tilannetietoisuus** ovat oleellisia teknologioita tekoälyn käytännön hyödyntämisen kannalta, ja ominaisuuksia on eritelty taulukossa 1.
3. **Luonnollinen kieli ja kognitio** muodostavat keskeisen teknologioiden joukon, jonka ominaisuuksia on luonnollisen kielen käsittelyyn keskittyen tarkasteltu taulukossa 2.
4. **Vuorovaikutus ihmisen kanssa** liittyy useisiin teknologioihin, kuten robotiikkaan, havainnointiin ja luonnolliseen kieleen. Tähän osa-alueeseen liittyvät ominaisuudet on käsitelty ko. teknologioiden yhteydessä.

5. **Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus** liittyvät tekoälyn käytännön soveltamiseen työelämän tehtävissä. Osa-alueeseen liittyviä ominaisuuksia käsitellään kunkin teknologian yhteydessä.
6. **Koneoppiminen** on keskeinen teknologiajoukko, joka käsitellään taulukossa 3. Toisaalta tarkastelemme myös malli- ja sääntöpohjaisten järjestelmien ominaisuuksia (taulukko 4).
7. **Järjestelmätaso ja systeemivaikutukset** ovat laajoja kokonaisuuksia, jotka vaikuttavat teknologioiden käyttöön ja valintaan. Niitä ei kuitenkaan esitetä erillisenä teknologiana, vaan vaikutuksia on kuvattu kunkin teknologian yhteydessä.
8. **Tekoälyn laskentaympäristöt, alustat ja palvelut ja ekosysteemit** mahdollistavat tekoälysovellusten kehittämisen. Käsittelemme laskentaympäristöjä osana fyysisiä laitteistoja (taulukko 5).
9. **Robotiikka ja koneautomaatio – tekoälyn fyysinen ulottuvuus** on myös käsitelty yhdessä laskentaympäristöjen kanssa osana fyysisiä laitteistoja (taulukko 5).
10. **Etiikka, moraali, regulaatio ja lainsäädäntö** ovat teknologian käyttöön liittyviä ulottuvuuksia, eikä niitä tässä yhteydessä käsitellä teknologian ominaisuuksina.

Keskeisiä tekoälyn hyödyntämisessä tarvittavia teknologioita on esitetty kuviossa 26.

Kuvio 26 Tässä raportissa tarkasteltavia tekoälyn hyödyntämisessä oleellisia teknologioita, joihin sisältyy sekä ohjelmistoteknologioita että erilaisia fyysisiä laitteistoja



4.2.1 Havainnoinnin ja tilannetietoisuuden teknologiat

Havainnointi ja tilannetietoisuus (ks. luku 2.2.2) voidaan nähdä katsantokannasta riippuen puhtaasti sovellusalueena tai omana teknologianaan. Tekoälyn käytännön soveltamisen kannalta havainnointiin liittyvät teknologiat ovat erittäin tärkeitä, minkä vuoksi niiden ominaisuuksia käsitellään tässä yhteydessä erillisenä teknologiakokonaisuutena. Erityisen oleellisia havainnointiin liittyvät kyvykkyydet ja sen rajoitteet ovat tapauksissa, joissa tekoäly on vuorovaikutuksessa ihmisten kanssa.

Havainnointiin käytettävät järjestelmät voivat olla koneoppimisen menetelmiin perustuvia, mutta käytössä on myös erilaisia mallipohjaisia järjestelmiä. Ohjelmistojen toteutuksen lisäksi on huomioitava myös fyysisen maailman ulottuvuus, kuten tilannetietoisuuden rakentamisessa tarvittavat sensorijärjestelmät. Taulukossa 1 on esitetty havainnointiin ja tilannetietoisuuteen liittyvien teknologioiden ominaisuuksia ja kehitysodotuksia.

Taulukko 1 Havainnointiin liittyvien teknologioiden ominaisuuksia sekä kehitysodotuksia

Havainnointi ja tilannetietoisuus		
Ominaisuus	**Tilanne nyt**	**Kehitysodotukset**
Oppivuus	• Konenäkösovellusten toiminta pääosin rajoitettua kulloiseenkin sovellusalueeseen	• Kuvantunnistus kehittyy yhä adaptiivisempaan suuntaan mahdollistaen laajemman hyödyntämisen
Suorituskyky	• Kuvantunnistus rajatuissa tapauksissa ihmisen kyvykkyyttä edellä • Havainnointijärjestelmien opetuksessa tarvittavan datan kerääminen ja hidastaa suorituskyvyn kehitystä • Sensorijärjestelmien suorituskyky ja luotettavuus eri olosuhteissa sovelluksesta riippuen rajoittavia tekijöitä	• Opetusdatan luonti simulaattoripohjaisesti nopeuttaa kehitystä • Suorituskyky tulee kasvamaan teknologian kehityksen myötä
Autonomia	• Järjestelmät eivät autonomisesti sopeudu uusiin tehtäviin	• Autonomia ja järjestelmien monikäyttöisyys kasvaa hitaasti
Avoimuus	• Avoimen lähdekoodin kirjastot tarjoavat kattavasti työkaluja	• Avoimet teknologiat säilyttävät asemansa
Riskit	• Havainnoinnin ongelmat vaihtuvissa olosuhteissa • Yksityisyysongelmat ihmisten havainnointiin liittyvissä sovelluksissa	• Käyttökohteiden lisääntyminen vaatii riskien systemaattisempaa hallintaa
Validointi	• Lopputuloksen validointi pääosin manuaalista	• Automatisoituja ja joukkoistusta hyödyntäviä keinoja tulosten validointiin

4.2.2 Luonnollisen kielen käsittely

Luonnollisen kielen käsittely (ks. luku 2.2.3) on kokoelma erilaisia lähestymistapoja ja teknologioita., joilla pyritään ymmärtämään sekä puhutun että kirjoitetun kielten merkityksiä ja vastaavasti tuottamaan puhetta ja tekstiä. Teknologian sovelluskohteet liittyvät mm. aineistojen käsittelyyn ja luokitteluun, puheentunnistukseen sekä konekääntämisen kaltaisiin sovelluksiin.

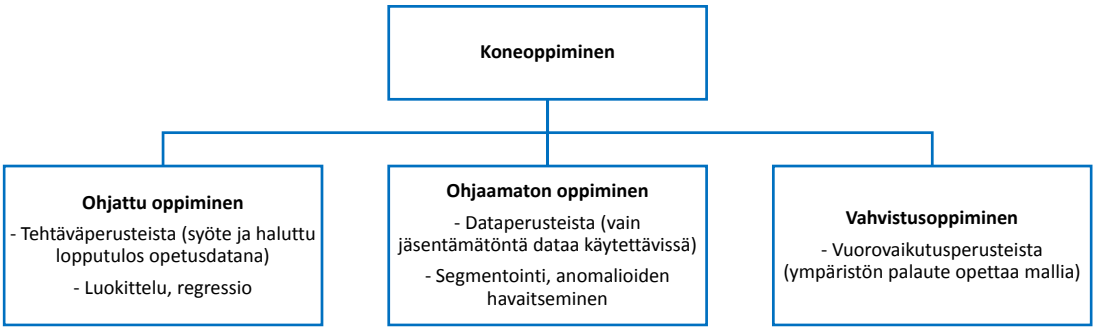
Taulukko 2 Luonnollisen kielen käsittelyssä käytettävien teknologioiden ominaisuuksia

Luonnollisen kielen käsittely		
Ominaisuus	**Tilanne nyt**	**Kehitysodotukset**
Oppivuus	• Käytössä sääntöpohjaisia, tilastollisia ja neuroverkkoihin perustuvia järjestelmiä, joista vain jälkimmäisissä oppivuutta	• Sääntöpohjaisista ja tilastollisista malleista siirrytään kiihtyvällä tahdilla neuroverkkopohjaisiin oppiviin sovelluksiin
Suorituskyky	• Useissa tehtävissä ihmisen tasolla tai edellä: puheentunnistus, kielenkääntäminen • Merkitysten ymmärtäminen haasteellista • Kielten välillä eroja suorituskyvyssä	• Suorituskyvyn kehitys mahdollistaa reaaliaikaisen kääntämisen useiden eri kielten välillä • Kielten väliset erot suorituskyvyssä pienenevät
Autonomia	• Oppiminen ei pääosin ole autonomista, vaan perustuu ennalta määriteltyihin reunaehtoihin	• Koneoppimiseen perustuvat menetelmät mahdollistavat autonomisemman toiminnan • Ihmisen ja tekoälyn välinen kommunikointi luonnollisella kielellä yleistyy nopeasti (käyttöliittymät, chatbotit)
Avoimuus	• Avoimen lähdekoodin kirjastot kärkitasoa useissa kielen käsittelyn tehtävissä • Rajatuissa ammattisovelluksissa sanastojen ja tilastollisten materiaalien omistus voi rajoittaa avoimuutta	• Ei odotettavissa merkittäviä muutoksia
Riskit	• Yksityisyysongelmat tekstin ja puheen tunnistuksessa • Merkitysongelmat käännöksissä, virheellinen opetusdata	• Riskit korostuvat teknologioiden yleistyessä ja kehittyessä
Validointi	• Validointi vaatii runsaasti manuaalista työtä, esim. Konekäännösten laadun tarkastaminen	• Joukkoistuksen keinoja tulosten validointiin

4.2.3 Koneoppiminen

Tekoälyn viimeaikaiset edistysaskeleet ovat monelta osin olleet koneoppimiseen menetelmiin perustuvia. Koneoppimisen perusmääritelmä, sekä yleisen tason esimerkkejä sen käytöstä on kuvattu luvussa 2.2.6. Tässä luvussa käsitellään tarkemmin koneoppimisen keskeisiä menetelmiä (kuvio 27), sekä niiden soveltuvuutta erilaisten ongelmatyyppien ratkaisemiseen (Jordan ja Mitchell, 2015). Yleisesti koneoppimista hyödynnetään tapauksissa, joissa ongelman kuvaaminen manuaalisesti määritetyillä säännöillä on liian monimutkaista, tai joissa tehtävän luonne muuttuu jatkuvasti siten, että oppivuus on välttämätöntä.

Kuvio 27 Koneoppimisen menetelmiä ja niiden käyttökohteita



Ohjattu oppiminen

Ohjattua oppimista (*supervised learning*) käytetään tilanteissa, joissa opetusdatassa on tiedossa sekä kunkin tapauksen syöte että haluttu lopputulos. Menetelmän tavoitteena on muodostaa opetusdatan perusteella malli, joka ennustaa uusien syötteiden perusteella uudet lopputulokset mahdollisimman tarkasti. Tyypillisiä ohjatun oppimisen sovelluskohteita ovat **luokittelu**- ja **regressio**-ongelmat. Luokittelulla tarkoitetaan tilanteita, joissa syötteitä sijoitetaan ennalta määriteltyihin luokkiin. Regressiotyyppisissä sovelluksissa puolestaan pyritään ennustamaan jatkuva-arvoisen muuttujan kehitystä. (MathWorks, 2018)

Enenevässä määrin ohjatun oppimisen sovelluksissa hyödynnetään neuroverkkoja (*Artificial neural network, ANN*), ja lähestymistavalla on saavutettu monet viimeaikaisista näyttävistä tuloksista tekoälyn soveltamisessa. Neuroverkot jäljittelevät karkeasti neuronien toimintaa ihmisaivoissa: luonnollisten hermoverkkojen tapaan ne muodostuvat lukuisista toisiinsa kytketyistä tietojenkäsittelyelementeistä. Varsinaiseen ihmisaivojen jäljittelyyn ei kuitenkaan yleensä pyritä, vaan verkon rakenne, koko ja kytkennät valitaan sovelluskohteen mukaisesti. Neuroverkkotyyppejä on runsaasti erilaisia, joista viime aikoina erityisesti useista neuronikerroksista koostuvat syvät neuroverkot (*deep neural networks*) ovat saaneet runsaasti huomiota. Kun tällaisten verkkojen opetuksessa käytetään runsaasti dataa, puhutaan myös syväoppimisesta (*deep learning*). (Lee ym., 2018)

Opetusdatan hyödyntäminen luokittelussa

Esimerkkinä luokittelutyyppisestä ongelmanratkaisusta voidaan käyttää sähköpostin roskapostisuodatukseen tarkoitettua ohjelmistoa. Ohjelmiston tavoitteena on luokitella viestit kahteen kategoriaan: turvallisiin ja roskaposteihin. Opetusdatana käytetään sähköpostiviestien kokoelmaa, jossa roskapostit on yksitellen merkitty. Datan perusteella voidaan opettaa ohjelma havaitsemaan tyypillisiä roskapostiviesteissä esiintyviä piirteitä, ja siten luokittelemaan ne oikeaan kategoriaan. Regressiotyyppisestä ongelmasta esimerkkejä puolestaan ovat mm. osakekurssien tai asuntojen hintakehityksen ennustaminen historiatietoon ja muihin parametreihin perustuen.

Kokonaisuutena ohjattu oppiminen on usein erittäin tehokas menetelmäjoukko, kun tarvittava tieto halutusta lopputuloksesta on saatavilla. Laadukkaan ja jäsennetyn opetusdatan hankkiminen voi kuitenkin aiheuttaa ongelmia käytännön toteutukselle, sillä opetusdatan koostaminen ja ohjatussa oppimisessa tarvittavien syöte-tulos-parien määritys voivat olla hyvin työläitä ja aikaa vieviä toimenpiteitä. Joissakin tapauksissa (esim. turvallisuuskriittiset sovellukset) opetusdataa ei-toivotuista poikkeustapauksista voi olla jopa mahdotonta kerätä.

Ohjaamaton oppiminen ja vahvistusoppiminen

Ohjaamattomassa oppimisessa (*unsupervised learning*, kuvio 27) ohjaamattomuus viittaa siihen, ettei ennustettavaa piirrettä ole etukäteen määritelty (vrt. ohjatun oppimisen syöte-lopputulos-parit). Ohjaamattomassa oppimisessa lähtökohta sen sijaan on dataperustainen: tarkoituksena on muodostaa malli, jonka avulla voidaan kuvata aineiston rakennetta esimerkiksi klusteroimalla sitä luokkiin tai havaitsemalla erilaisia anomalioita. Toisin sanoen menetelmää hyödynnetään, kun jonkin piirteen ennustamisen sijaan ollaan erityisesti kiinnostuneita aineiston rakenteesta ja siinä esiintyvistä säännönmukaisuuksista. (Lee ym., 2018)

Taulukko 3 Koneoppimisen ominaisuuksia ja kehitysodotuksia

Koneoppiminen		
Ominaisuus	**Tilanne nyt**	**Kehitysodotukset**
Oppivuus	• Oppivuus rajoittunutta tarkoin määriteltyyn sovelluskohteeseen	• Oppivuuden lisäksi kiinnostus luovia ominaisuuksia kohtaan kasvaa, ts. miten tekoäly opitun perusteella luo uutta
Suorituskyky	• Käytännön sovellukset usein laskentaintensiivisiä, jolloin suorituskyky riippuu laskentakapasiteetin määrästä (Sun et al., 2017) • Ohjattu oppiminen vahvasti riippuvaista opetusdatan laadusta ja määrästä	• Opetusdatan määrät kasvavat vaatien samalla jatkuvaa edistystä laskentaympäristöihin liittyen
Autonomia	• Sopeutuu heikosti odottamattomiin tilanteisiin tai vaihtuvaan käyttöympäristöön	• Autonomia kasvaa, kun kehitetään paremmin epävarmuuksia sietäviä järjestelmiä (Alexandre, 2016). Nopeita kehitysharppauksia ei kuitenkaan ole odotettavissa.
Avoimuus	• Tarjolla kaupallisia ja avoimien lähdekoodin teknologioita • Algoritmien avoimuus rajoitettua	• Sekä avoimet että kaupalliset teknologiat säilyttänevät asemansa
Riskit	• Opetusdatan saatavuus ja laatu asettavat haasteita kehitykselle • Päättelyn läpinäkymättömyys (neuroverkkosovellukset)	• Koneoppimisen käyttöönotto uusissa sovelluskohteissa tuo mukanaan uudenlaisia riskejä
Validointi	• Tarvitaan paitsi oppimisen lopputulosten validointia, myös datan laatuun liittyvää validointia • Standardointi ei pysy teknologiakehityksen vauhdissa: teknologian kehittäjällä suuri vastuu validoinnista	• Validointi yhä haastavampaa järjestelmien monimutkaistuessa ja datamäärien kasvaessa

Ohjatun ja ohjaamattoman oppimisen välimaastoon sijoittuu lisäksi vahvistusoppimisena (*reinforcement learning*) tunnettu menetelmä, jossa algoritmi optimoi toimintaansa ympäristön antaman palautteen perusteella siten, että positiivisten vasteiden määrä kasvaa. Algoritmia siis ”palkitaan” hyvistä valinnoista ja siten vahvistetaan haluttua toimintaa. (Mnih ym., 2015)

Taulukossa 3 on esitetty koneoppimismenetelmien ominaisuuksia ja niihin liittyviä kehitysodotuksia. Ominaisuuksien määrittely on esitetty edellä luvussa 4.1.

4.2.4 Malli- ja sääntöpohjaiset järjestelmät

Koneoppimiseen perustuvien menetelmien lisäksi tekoälyteknologioiksi voidaan lukea erilaisia malli- ja sääntöpohjaisia järjestelmiä, vrt. luku 2.2.7.

Innostus tekoälyä kohtaan on ollut sen historian aikana vaiheittain nousevaa ja laskevaa. 1980-luvulla erityisesti sääntöpohjaiset asiantuntijajärjestelmät olivat intensiivisen tutkimuksen kohteina. Asiantuntijajärjestelmän tarkoituksena on mukailla ihmisen päätöksentekoa määrätyssä rajatussa sovelluskohteessa.

Käytännössä asiantuntijajärjestelmä voi olla toteutettu esimerkiksi päätöspuun avulla. Päätöspuulla tarkoitetaan mallia, joka luokittelee syötteen määrätyn päätösrakenteen mukaisesti. Päätöspuut voidaan sijoittaa toteutustavasta riippuen koneoppimisen alaluokaksi tai omaksi teknologiaperheekseen. Tässä yhteydessä on haluttu korostaa, etteivät tekoälynä pidettävät teknologiat välttämättä rajoitu pelkästään esimerkiksi neuroverkkopohjaisiin koneoppimissovelluksiin, vaan sovelluskohteesta riippuen esimerkiksi sääntöperustainen lähestymistapa voi olla tehokkaampi.

Taulukko 4 Malli- ja sääntöpohjaisten järjestelmien ominaisuuksia

Mallipohjaiset järjestelmät		
Ominaisuus	**Tilanne nyt**	**Kehitysodotukset**
Oppivuus	• Mallipohjaiset järjestelmät perustuvat pääasiassa manuaalisesti luotuihin sääntöihin, oppivuus ei olennaista	• Mallipohjaisuuden ja koneoppimismenetelmien yhdistäminen avaa uusia mahdollisuuksia
Suorituskyky	• Sovellukset suorituskykyisiä tarkoin rajatuissa tehtävissä • Vaativat vähemmän laskentakapasiteettia kuin datapohjaiset ratkaisut	• Ei odotettavissa merkittävää muutosta
Autonomia	• Voivat toimia autonomisesti rajatussa tehtävässään, mutta eivät sovellu muutoksiin	• Ei odotettavissa merkittävää muutosta
Avoimuus	• Pääosin suljettuja, yritysten sisäisiä järjestelmiä • Geneeriset mallit yleensä avoimia	• Ei odotettavissa merkittävää muutosta
Riskit	• Virheet tai rajoitteet mallissa mahd. vaikeita havaita	• Ei odotettavissa merkittävää muutosta
Validointi	• Voidaan käyttää esim. mallintarkastusmenetelmiä • Muutostilanteissa validointi voi olla ongelmallista	• Ei odotettavissa merkittävää muutosta

4.3 Fyysiset laitteistot tekoälyn mahdollistajina

4.3.1 Laskentaympäristöt

Pitkään suhteellisen tasaisena jatkunut laskentakapasiteetin kasvu on ollut osaltaan merkittävästi vauhdittamassa tekoälyteknologioiden nopeaa esiinmarssia, ks. myös 2.2.8. Prosessointitehon kustannus on laskenut jo vuosikymmenten ajan ns. Mooren lakia mukaillen, mutta viime aikoina kyseinen kehitys on hidastunut. Samalla kasvavat datamäärät ja yleistyvät resurssi-intensiiviset tekoälyn laskentamenetelmät vaativat yhä enemmän laskentakapasiteettia. Tämä aiheuttaa tarvetta erilaisille tekoälytarkoituksiin optimoiduille laitteistoille.

4.3.2 Robotiikka ja vuorovaikutus fyysisen maailman kanssa

Tekoälyn mahdollisuudet eivät rajoitu pelkästään digitaaliseen maailmaan, vaan tekoälyjärjestelmien vuorovaikutus fyysisen maailman kanssa korostuu, ks. 2.2.9. Vuorovaikutuksessa ihmisten ja ympäröivän maailman kanssa tarvitaan uudenlaisia käyttöliittymäratkaisuja ja robotiikan sovelluksia. Sovelluskohteita on mm. kaikilla liikenteen aloilla ja teollisuudessa, kun ajoneuvoista, työkoneista ja teollisuuden järjestelmistä kehitetään yhä autonomisempia. Toisaalta robotiikassa myös työskentely ihmisen kanssa korostuu, kun kehitetään yhteistyörobotteja esim. terveydenhuoltoalan tehtäviin.

Taulukko 5 Fyysisiin laitteistoihin liittyviä ominaisuuksia ja kehitysehdotuksia

Fyysiset laitteistot - laskentaympäristöt ja robotiikka		
Ominaisuus	**Tilanne nyt**	**Kehitysodotukset**
Oppivuus	• Yhteistyörobotiikka mahdollistaa yksinkertaisten tehtävien opetuksen	• Robotiikan oppivuus lisääntyy
Suorituskyky	• Perinteinen laskentakapasiteetin kasvuun perustuva suorituskyvyn kasvu hidastumassa	• Laskentaympäristöjen suorituskyky AI-sovelluksissa jatkaa nopeaa kasvuaan AI-käyttöön optimoinnin myötä • Teollisuudessa yksittäisten robottien suorituskykyä suuremmat hyödyt saavutetaan tehokkaalla yhdistelmällä useiden robottien välistä yhteistyötä sekä yhteistyötä ihmisen kanssa
Autonomia	• Autonomisuus alussa testiasteella	• Autonomian taso kasvaa eri sovellusalueilla, esim. liikenne, teollisuuden sovellukset • Autonomisten järjestelmien kasvuodotukset suuria • Robottien interaktio ihmisten kanssa lisääntyy
Avoimuus	• Avoimen lähdekoodin ympäristöjä saatavilla, mutta useat teknologioista yritysten omistuksessa • Robottivalmistajilla omia, patentoituja ratkaisujaan	• Suuret toimijat pyrkivät luomaan teknologioistaan ”de facto -standardeja”, jolloin avoimuus voi vaarantua
Riskit	• Turvallisuusongelmat toimittaessa samassa ympäristössä ihmisten kanssa • Kyberturvallisuus, robotiikan käyttö vahingollisissa tarkoituksissa • Käyttötarkoituksen muuntaminen	• Robotiikan ja autonomisten laitteiden käyttö entistä avoimemmassa ympäristössä tuo uudenlaisia riskejä
Validointi	• Perustuu perinteisiin validointimenetelmiin	• Tarvitaan uusia validointimenetelmiä, jotka tukevat järjestelmien kokonaisvaltaista arviointia

LIITE 1
TEKOÄLY – NELJÄ TÄYDENTÄVÄÄ NÄKÖKULMAA

Tässä liitteessä tarkastellaan lyhyesti tekoälyä raportissa käytettyä kymmenen osaamisalueen näkökulmaa täydentävistä kulmista:

- kehitysasteiden mukaan,
- Darpan ”kolmen aallon” näkökulmasta,
- Stanfordin AI 100 -paneelin kuumat tutkimusaiheet,
- tekoälyn koulukunnat ja menetelmät,
- tulevaisuuden trendit tekoälyssä.

Tekoälyn kehitysasteet: kapea, yleinen ja super-tekoäly

Tekoäly jaetaan perinteisesti heikkoon (*weak AI*) ja vahvaan tekoälyyn (*strong AI*) (mm. Russell ja Norvig, 2014). Kapean tekoälyn määritelmä on, että kone toimii järkevästi tai älykkäästi, kun taas vahva tekoäly edellyttää koneelta älykästä, tyypillisesti ihmisen kaltaista, ajattelua ja jonkinasteista tietoisuutta.

Vastaavasti puhutaan kapeasta ja yleisestä tekoälystä (*narrow vs. general* tai *broad AI*), jolloin määrittely perustuu tarkasteltavan tekoälyn kyvykkyyksien laajuuteen ja hyödyllisyyteen erilaisissa tilanteissa tai sovelluksissa (Muehlhauser ja Salamon, 2012). Kapea ja heikko ja vastaavasti vahva ja yleinen tekoäly ovat käsitteinä lähellä toisiaan ja niitä käytetään jopa synonyymeinä, vaikka tarkkaan ottaen niiden välillä on tietty ero. Käytämme tässä raportissa termiä kapea tekoäly, koska ”heikko tekoäly” ei tee oikeutta tekoälyn suorituskyvylle niissä tehtävissä, joihin sitä sovelletaan.

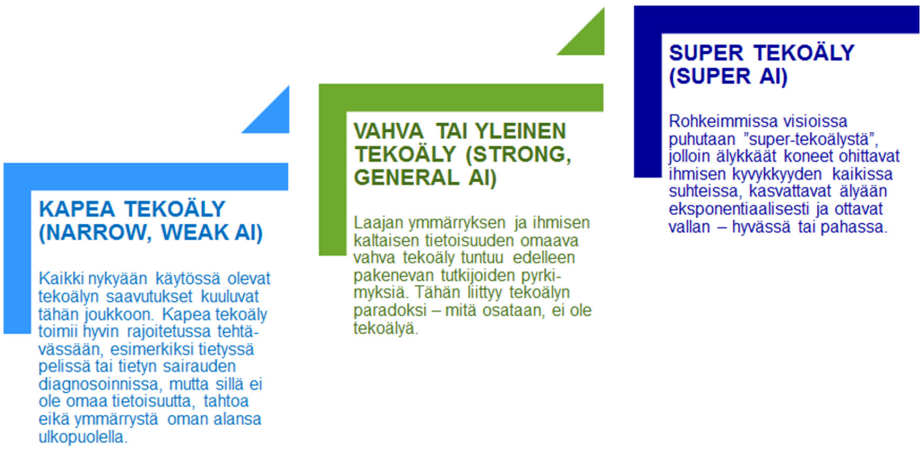
Usein puhutaan myös ”super tekoälystä” (*Super AI*, *superhuman AI*), jossa tekoäly ylittää ihmisen älyllisen kyvykkyyden kaikissa suhteissa. Tämän katsotaan olevan vääjäämätön seuraus teknologian kiihtyvästä kehityksestä, esim. Mooren laista. Tässä skenaariossa älykkäät ohjelmistot tai agentit alkavat kehittää itseään yhä nopeammin kiihtyvän ketjureaktion tapaan ja ohittavat ja ylittävät inhimillisen suorituskyvyn. Kehitys voidaan nähdä positiivisena tai dystopisena. (Vinge, 1993; Muehlhauser ja Salamon, 2012; Bostrom, 2014.)

On syytä huomata ero Ray Kurzweilin (Kurzweil, 2010) propagoimaan singulariteetti-hypoteesiin, joka esittää, että teknologian kehitys tuo ihmisille uusia kykyjä, joiden ansiosta syntyy uusi ”posthuman” ihmisrotu. Tämä ”uusi ihminen” on vapaa monista biologisista rajoitteista, kuten sairauksista ja jopa kuolemasta, ja hänellä on ylivertaisia kognitiivisia kykyjä meihin verrattuna. Kurzweil ajoittaa singulariteetin vuoteen 2045.

Tutkijoiden kesken on laaja yksimielisyys siitä, että kapea tekoäly on taso, johon tekoäly tutkimuksessa ja soveltamisessa on tähän mennessä päästy; yleistä tai vahvaa tekoälyä ei ole kyetty toteuttamaan (Dicson, 2017). Samalla on syytä huomata, että tekoäly on ylittänyt ihmisen suorituskyvyn useissa vaativissa ja aiemmin inhimillistä älykkyyttä osoittavina pidetyissä tehtävissä kuten kasvojen tunnistamisessa, kielen kääntämisessä tai visailuissa.

Tekoälyn kehitysasteet on esitetty tiivistetysti kuviossa L1.

Kuvio L1 Tekoälyn kehitysasteet



Tekoälyn kolme aaltoa Darpan mukaan

On syytä vielä esittää Darpan (Yhdysvaltain puolustusministeriön alainen tutkimusrahoituslaitos) näkemys tekoälyn tasoista, koska laitos on paitsi merkittävä tutkimuksen rahoittaja, myös ajatusjohtaja.

DARPA kuvaa tekoälyn kehitystä kolmen aallon avulla, jotka ovat ”kuvaa, luokittele ja selitä” (*describe*, *categorize and explain*) (Laundsbury, 2017). Aaltoihin liittyvät seuraavat ominaisuudet:

– aalto, ”kuvaileva”, käsin muokattu tietämys, s.o. sääntöpohjaiset ihmisen rakentamat mallit, soveltuu kapeisiin selkeisiin tehtäviin, vakaaseen ympäristöön, ei oppiva, mutta jonkin verran päättelykykyä
– 2. aalto, ”luokitteleva”, tilastollinen oppiminen, koneoppiminen, neuroverkot, suuret data määrät, kapea-alainen, ei päättelykykyä
– 3. tuleva aalto, ”selittävä”, kontekstuaalinen sopeutuminen, havainnoi, oppii, päättelee ja abstrahoi.

Stanfordin AI100 -paneelin näkemys

Stanfordin yliopiston käynnistämän merkittävän tutkimushankkeen ”Artificial intelligence and life in 2030 – one hundred year study on AI” vuosiraportti (Report of the study panel 2015 ja 2017) nostaa esiin seuraavat tutkimustrendit, joita voidaan pitää tekoälyn ”kuumina” aiheina nyt.

– Koneoppimisen menetelmien soveltaminen massiivisen data-aineiston analyysissä.
– Syväoppiminen erityisesti kuva-analyysissä ja muussa havainnoinnissa sekä luonnollisen kielen käsittelyssä (*deep learning neural networks*). Tähän liittyen kehitetään neuroverkkolaskentaan perinteisiä prosessoreita paremmin soveltuvien ns. neuromorfisten tietokoneiden kehitys.
– Vahvistettu oppiminen (*reinforcement learning*) päätöksenteon tueksi todellisissa ongelmissa.
– Robotiikka, autonomiset ajoneuvot, esineiden käsittely.

– Kuva-analyysi ja konenäkö, jotka ovat hyötyneet syväoppivien neuroverkkojen kehityksestä.
– Luonnollisen kielen kääntäminen, erityisesti puheen tunnistus ja kielen kääntäminen.
– Ihmisen ja tietokoneen yhteistyö ja joukkoistaminen, jossa tehtävä jaetaan ihmisille ja tietokoneille niin, että kummankin vahvuudet päästään hyödyntämään.
– Peliteorian soveltaminen kannustimiin ja muihin sosiaalisiin ja liike-elämän kysymyksiin.
– Tekoäly nähdään keinona yhdistää ja hyödyntää tehokkaasti esineiden internetin (IoT) tuottama valtava ja sekalainen informaatio.

Tässä yhteydessä korostetaan, että aiheen nykyinen suosio ei välttämättä korreloi sen pitkän aikavälin merkittävyyden kanssa, vaan kertoo siitä, että aiheessa on saavutettu viime vuosina tai odotetaan lähivuosina nopeaa edistystä.

Keskeiset tekoälymenetelmät on pyritty luokittelemaan taulukossa L1.

Keskeiset tekoälyn koulukunnat ja menetelmät

Tekoäly- ja kognitiotutkimus voidaan jakaa koulukukuntiin ja näihin voidaan liittää tiettyjä teknologioita ja menetelmiä. Perinteisesti on tunnettu kaksi pääkoulukuntaa: konnektionismi (connectionistic) ja symbolinen tekoäly.

Konnektionismi, jonka mukaan mielen ilmiöt (ajattelu, tunteet, tietoisuus) syntyvät yksinkertaisten ja tyypillisesti keskenään samankaltaisten toisiinsa kytkeytyneiden yksiköiden toiminnan tuloksena. Käytännössä nämä yksiköt ovat neuroneja ja synapseja; tekoälystä puhuttaessa (keinotekoisia) neuroverkkoja. Joskus käytetään myös termiä computational intelligence suunnilleen samassa merkityksessä. Tutkii – ja ratkaisee – tehokkaasti rajattuja ongelmia: shakkipeli, kasvojen tunnistus. Konnektionismi voidaan liittää datapohjaisuuteen.

Symbolinen tekoäly (*symbolic AI*) käsittää tekoälymenetelmät, joissa tietoa käsitellään ihmisen ymmärtämien korkeamman tason symbolien avulla, esimerkiksi logiikalla tai päättelysäännöillä. Joskus tätä kutsutaan termillä GOFAI (*Good Old-Fashioned Artificial Intelligence*). Tämä suuntaus hallitsi tutkimusta 1950-luvulta 1980-luvun lopulle. Ontologiat ovat yksi tapa esittää symbolista tietoa, ja semanttiset menetelmät ovat välineitä symbolisen tiedon käsittelyyn. Tutkii laajempia ja abstrakteja tutkimuskysymyksiä (*general AI*). Tämä koulukunta voidaan liittää mallipohjaisuuteen. Näitä kahta on pidetty tekoälyn pääkoulukuntina, esim. Hoffman (1998).

Kolmantena koulukuntana voidaan mainita Embodied Intelligence eli ”Kehollistettu” tekoäly, joka korostaa liikkumisen ja älyn yhteyttä. Ajatuksen kiteyttää seuraava lainaus Lewis Wolpertilta: ”Why do plants not have brains? The answer is actually quite simple: they don’t have to move.” Älykkään toiminnan ajatellaan syntyvän emergentisti anturi–toimilaite-silmukoista, kun organismi tai laite on vuorovaikutuksessa ympäristönsä kanssa. Kehollisen älykkyyden tutkimuksen mallina ovat usein biologiset organismit ja tutkimusvälineinä liikkuvat mekanismit tai robotit.

Taulukko L1 jäsentää tekoälyn koulukunnat ja menetelmät. On syytä huomata, että tämä on karkea yksinkertaistus eikä yritäkään luetella kaikkia varteenotettavia tekoälymenetelmiä.

Taulukko L1 Tekoälyn koulukunnat ja menetelmät

<table>
<tr><th colspan="5">Tekoälyn koulukunnat ja menetelmät
(kaikkia menetelmiä ei luetella, vain keskeiset esimerkit)</th></tr>
<tr><th colspan="3">Datapohjaiset menetelmät
~ ”Connectionistic” –koulukunta</th><th>Symbolinen tekoäly
~Klassinen koulukunta</th><th>”Kehollistettu tekoäly”
~Embodied AI</th></tr>
<tr><td><u>Ohjattu oppiminen</u></td><td><u>Ohjaamaton oppiminen</u></td><td><u>Muut</u></td><td>Semantiikka, ontologiat</td><td rowspan="6">Korostaa liikkumisen ja ”älyn” yhteyttä</td></tr>
<tr><td>Lineaarinen regressio</td><td>PCA, LCA</td><td rowspan="5">Geneettiset algoritmit</td><td>Edellisiin perustuva logiikka</td></tr>
<tr><td>Neuroverkot</td><td>Neuroverkot</td><td>Haku (search)</td></tr>
<tr><td>Tukivektorikoneet SVM</td><td>SOM</td><td>Suunnittelu (Planning)</td></tr>
<tr><td>Logistinen regresessio</td><td>Poikkeavuuksien havaitseminen</td><td rowspan="2">Päätöspuut, asiantuntijajärjestelmät</td></tr>
<tr><td>Lineaarinen erotteluanalyysi LDA</td><td>GAN-verkot</td></tr>
</table>

Tulevaisuuden trendit tekoälyssä

Datapohjaiset menetelmät ovat osoittautuneet hyvin suorituskykyisiksi, ja sen vuoksi on odotettavissa, että niiden tutkimiseen, kehittämiseen ja soveltamiseen panostetaan myös näköpiirissä olevassa tulevaisuudessa. Samalla tämä tarkoittaa pienempää huomiota aiemmin valtavirtaa edustaneille tutkimussuunnille (”klassiselle tekoälylle”), kuten tietämys- ja päättelymenetelmille, jossa pyrittiin mallintamaan reaalimaailman ilmiöitä. Osittain syynä vähentyneelle kiinnostukselle mallipohjaisia menetelmiä kohtaan on törmääminen niiden rajoitteisiin, esimerkiksi mallien vaatiman pelkistämisen ja reaalimaailman monimutkaisuuden välillä.

Olisi kuitenkin lyhytnäköistä hylätä muita menetelmiä syvien neuroverkkojen ja datapohjaisuuden viime vuosien menestyksen vuoksi. Myös niiden puutteet ja soveltamisen rajat tulevat esiin lähivuosina. Tekoälyteknologioihin liittyvä seuraava trendi tulee mahdollisesti liittymään ihmisen ja koneen tai järjestelmän välisen yhteistoiminnan parantamiseen sekä esineiden internetin (IoT) tarjoamiin mahdollisuuksia. (AI100, 2016).

LIITE 2 JULKAISUANALYYSIN NÄKÖKULMAT JA MENETELMÄT

Jokaisesta kymmenestä tutkimussuunnasta ja yleisesti rajatusta tekoäly-teemasta käytiin läpi Scopus-tietokannasta saatujen julkaisumäärien kehitys v. 2000–2017 (v. 2017 tilanne: 04/2018) sekä keskeiset kyseisen tutkimussuunnan aihepiirit ja teemojen tutkimusintensiteetin muutos (julkaisumäärien keskimääräinen vuosimuutos) v. 2008–2017.

Lisäksi seuraavista tutkimussuunnista käytiin läpi keskeiset maat, maiden tutkimusintensiteetin muutos v. 2015–2018 ja painotukset yleisimmissä tutkimusaiheissa (maiden vertailu). Johtavien maiden lisäksi tarkasteluun otettiin Pohjoismaat ml. Suomi:

- Data-analyysi
- Havainnointi ja tilannetietoisuus
- Luonnollinen kieli ja kognitio
- Koneoppiminen
- Robotiikka ja koneautomaatio – tekoälyn fyysinen ulottuvuus
- Etiikka, moraali, regulaatio ja lainsäädäntö

Lisäksi selvitettiin myös keskeiset tutkimusorganisaatiot ja organisaatioiden tutkimusintensiteetin muutos v. 2015–2018 ja painotukset yleisimmissä tutkimusaiheissa seuraavissa tutkimussuunnissa:

- Koneoppiminen
- Tekoäly yleisesti

Eri maiden ja tutkimuskentän toimijoiden tutkimusprofiilien eroja vertailtiin tässä selvityksessä vain yleisellä tasolla. Tarkastelussa olivat julkaisujen avaintermit sellaisenaan ilman termien sisältöihin liittyviä tulkintoja, termien yhdistämistä tai muuulla tavoin tapahtuvaa muokkausta.

Julkaisujen tieteellisen vaikuttavuuden arviointi

Tieteellisen tutkimuksen vaikuttavuutta voidaan arvioida useista näkökulmista. Tässä selvityksessä on käytetty lähdeaineistona tutkimuksen tuottamia julkaisuja. Perinteisten bibliometristen mittarien – julkaisumäärien, viittausmäärien ja lehtien ns. impact factorien – rinnalle on viime vuosina noussut lukuisia niitä täydentäviä ja täsmentäviä tieteellisen vaikuttavuuden arvioinnin välineitä.

Mittarit voidaan jakaa karkeasti kolmeen ryhmään: a) lähdeviittausten lukumääriin, b) käytettyyn julkaisukanavaan ja nykyisin myös c) julkaisun käyttöön/näkyvyyteen/jakamiseen perustuviin mittareihin. Näiden lisäksi tieteellistä vaikuttavuutta tarkastellaan usein tutkimusyhteistyön näkökulmasta: yhteistyön on tutkimuksissa todettu mm. lisäävän julkaisujen viittausmääriä.[13]

13 Katso esim. Bornmann, L. (2017), Is collaboration among scientists related to the citation impact of papers because their quality increases with collaboration? Journal of Association for Information Science and Technology. Vol. 68, Issue 4, pp. 1036–1047; Nomalera, Ö., Frenkena, K., Heimeriks, G. (2013), Do more distant collaborations have more citation impact. Journal of Informetrics, Issue 7, pp. 966–971.

Mikään yksittäinen mittari ei silti itsessään tarjoa kattavaa näkemystä tutkimuksen tieteellisestä vaikuttavuudesta, vaan eri mittareita on aina syytä tarkastella rinnakkain. Erityisen tärkeää on huomioida myös julkaisutoiminnan volyymi silloin, kun tarkastellaan julkaisujen vaikuttavuutta. Pienet julkaisumäärät tuottavat suurta vaihtelua tieteellisen vaikuttavuuden arvioinnin mittareihin.

Tutkimustulosten vaikuttavuuden arvioinnin erityisenä haasteena voi pitää viivettä, joka syntyy julkaisun ilmestymisen ja julkaisun keräämien lähdeviittausten välille. Koska tämän selvityksen eräs keskeinen lähtökohta oli tarjota mahdollisimman ajantasainen kokonaiskuva – tilannekuva – tekoälystä, niin painopiste julkaisuaineistoissa oli myös vaikuttavuusselvitysten osalta uudemmassa aineistossa. Tällöin myös lähdeviittausten keräämisen aikaikkunaa jouduttiin joissakin kohdin supistamaan.

Tutkimuksen tieteellistä vaikuttavuutta tarkasteltiin selvityksessä kuudella mittarilla:

1. Suhteutettu viittausten lukumäärä (tutkimusala, julkaisuvuosi ja julkaisutyyppi huomioitu)
2. Julkaisujen määrä
3. Käytettyjen julkaisukanavien merkittävyys (top 10 % -lehtien/julkaisukanavien osuus)
4. Kansainvälinen julkaisuyhteistyö
5. Yritys-yliopisto -yhteistyö
6. Normalisoitu julkaisutietojen katselumäärä (lähteenä Scopus-tietokanta)

Vaikka selvityksessä on painotettu nimenomaan suhteutettua viittausten lukumäärää, niin kaikkia biblio- ja muun metriikan mittareita on syytä tarkastella tässä yhteydessä rinnakkain. Näin erityisesti, koska selvityksen tarpeista johtuen painopiste oli uusimmassa julkaisuaineistossa.

Tutkimussuuntien analysointi

Tutkimussuuntia tarkasteltiin yllä mainituista näkökulmista taulukon L2 mukaisesti.

Julkaisukatsaus on kokonaisuudessaan luettavissa verkossa osoitteessa https://tietokaytoon.fi/documents/1927382/2158283/Teko%C3%A4lyn+huippuosaamisen+kartoitus/0f93a62e-f97a-464d-9946-ba3953914d04. Kyseinen laaja katsaus sisältää kaikkien kymmenen tutkimussuuntauksen julkaisuselvitysten lisäksi myös lyhyen katsauksen alan patentointiin. Tässä valtioneuvoston VN TEAS Tekoälyn kokonaiskuva -väliraportissa käsitellään tekoälyn huippuosaamista ja tutkimuksen vaikuttavuutta vain yhtenä kokonaisuutena. Yhteenvedossa esitetyt johtopäätökset pohjautuvat kuitenkin kaikkien kyseisten kymmenen tutkimussuunnan julkaisuanalyyseihin.

Taulukko L2

Tutkimussuunta	Julkaisu-kehitys 2000–2017	Aiheet: volyymit ja muutokset 2008–2017	Maat: 1) volyymit ja muutokset 2003–2017 2) profiilivertailu 2015–2018	Tutkimusorganisaatiot: 1) volyymit ja muutokset 2003–2017 2) profiilivertailu 2015–2018	Julkaisujen tieteellisen vaikuttavuuden arviointi: 1) maat ja 2) tutkimus-organisaatiot
Tekoäly yleisesti tulkittuna	x	x	x	x	x
Data-analyysi	x	x	x		x[16]
Havainnointi ja tilannetietoisuus	x	x	x		x
Luonnollinen kieli ja kognitio	x	x	x		x
Vuorovaikutus ihmisen kanssa	x	x	x		
Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus	x	x			
Koneoppiminen	x	x	x	x	x
Järjestelmätaso ja systeemi-vaikutukset	x	x			
Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit	x	x			
Robotiikka ja koneautomaatio – tekoälyn fyysinen ulottuvuus	x	x	x		x
Etiikka, moraali, regulaatio ja lainsäädäntö	x	x	x		x

16 Data-analyysi -teeman osalta tieteellisen vaikuttavuuden arviointi jaettiin kahtia: 1) big data, 2) data mining.

LÄHTEITÄ JA TAUSTA-AINEISTOJA

AAAI2018 (2018). https://aaai.org/Conferences/AAAI-18/aaai18call (haettu 12.4.2018).

AI100 (2016). One Hundred Year Study on Artificial Intelligence (AI100). Stanford University. https://ai100.stanford.edu (haettu 3.12.2017).

Ailisto, H., Helaakoski, H., Dufva, M. ja Tuikka, T. (2017). Tuottoa ja tehokkuutta Suomeen tekoälyllä. VTT Policy Brief 1/2017.

Alexandre, F. (2016). Autonomous Machine Learning. ERCIM News 107.

Alpcan, T., Sarah, M., Erfani, S. ja Leckie, C. (2017). Toward the Starting Line: A Systems Engineering Approach to Strong AI. https://arxiv.org/abs/1707.09095 (haettu 13.4.2018).

Anderson, M. ja Anderson, S. (toim.) (2011). Machine ethics. Cambridge University Press.

Arntz, M., Gregory, T. ja Zierahn, U. (2016). The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. OECD Social, Employment and Migration Working Papers, No. 189, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jlz9h56dvq7-en

Asilomar AI Principles. https://futureoflife.org/ai-principles (haettu 28.5.2018).

Ballard, D. ja Brown, C. (1982). *Computer Vision*. Prentice Hall.

Bishop, P. (2006). Pattern Recognition and Machine Learning. Springer Verlag.

Blind, K. (2013). The Impact of Standardization and Standards on Innovation. Nesta Working Paper 13/15. November 2013. www.nesta.org.uk/wp13-15

Bostrom, N. (2014). Superintelligence, Paths, Dangers, Strategies. Oxford University Press, Oxford, UK.

Bostrom, N. (2017). Strategic Implications of Openness in AI Development. Future of Humanity Institute, University of Oxford, United Kingdom, as a part of Global Policy (2017).

Colton, S., Pease, A., Corneli, J., Cook, M., Hepworth, R. ja Ventura, D. (2015). Stakeholder Groups in Computational Creativity Research and Practice, teoksessa Besold, T., Schorlemmer, M. ja Smaill, A. (eds.) Computational Creativity Research: Towards Creative Machines. Atlantis Thinking Machines, vol 7. Atlantis Press, Paris.

Correll, N. (2016). Introduction to Autonomous Robots. v1.7, October 6, 2016 Magellan Scientific.

Craig, J. (2005). Introduction to Robotics – Mechanics and Control. 3rd edition. Pearson Prentice Hall, New Jersey, USA.

Crevier, D. (1993). AI: The Tumultuous Search for Artificial Intelligence, New York, Basic Books.

Dawson, M. (2013). Mind, Body, World – Foundations of Cognitive Science. AU Press, Athabasca University, Edmonton, Kanada. http://www.aupress.ca/books/120227/ebook/99Z_Dawson_2013-Mind_Body_World.pdf

DESI (2018). https://ec.europa.eu/digital-single-market/en/desi (haettu 9.4.2018).

Dicson, B. (2017). https://bdtechtalks.com/2017/05/12/what-is-narrow-general-and-super-artificial-intelligence (haettu 13.4.2018).

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Hum. Factors, 37 (1) (1995), pp. 32–64.

Erman, L. ja Lessner, V. (1977). System engineering techniques for artificial intelligence systems. Carnegie Mellon University Tech report.

EU (2018). Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions – Artificial Intelligence for Europe. https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe (haettu 10.5.2018).

Future of Life (2017). Ethics of AI. https://www.cs.ox.ac.uk/efai/towards-a-code-of-ethics-for-artificial-intelligence/what-is-ethics-anyway/ethics-and-law

Giles, M. (2018). Intelligent Machines: The GANfather: The man who's given machines the gift of imagination. MIT Technology Review, Vol. 121, No. 2, pp. 49–53.

Grace, K., Salvatier, J., Dafoe, A., Zhang, B. ja Evans, O. (2017). When Will AI Exceed Human Performance? Evidence from AI Experts. https://arxiv.org/pdf/1705.08807.pdf

Hoffman, A. (1998). Paradigms of Artificial Intelligence – A Methodological and Computational Analysis. Springer Verlag.

Hyvönen, E. (2018). Semanttinen web. Linkitetyn avoimen datan käsikirja. Gaudeamus. https://www.gaudeamus.fi/semanttinen-web

The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems. Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems, Version 1. IEEE, 2016. http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html

IEEE Std 1012-2016 (Revision of IEEE Std 1012-2012/ Incorporates IEEE Std 1012-2016/ Cor1-2017) – IEEE Standard for System, Software, and Hardware Verification and Validation.

IJCAI (2016). Special track on AI & The Web. http://ijcai-16.org/index.php/welcome/view/web_track (haettu 24.5.2018).

IMD (2018). https://www.imd.org/globalassets/wcc/docs/release-2017/world_digital_competitiveness_yearbook_2017.pdf (haettu 22.3.2018).

Jordan, M. I. ja Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255–260.

Judd, C. ja McCleland, G. (1989). Data Analysis. Harcourt Brace Jovanovich.

Kurzweil, R. (2010). The Singularity is Near. Duckworth & Co.

Lambda (2017). http://lambda-architecture.net (haettu 21.5.2018).

Laundsbury, J. (2017). The DARPA Perspective on AI. https://www.darpa.mil/attachments/AIFull.pdf (haettu 22.3.2018).

Lecun, Y., Bengio, Y. ja Hinton, G. (2015). Deep learning. Nature, 521(7553), pp. 436–444.

Lee, J. H., Shin, J. ja Realff, M. J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. Computers & Chemical Engineering. In Press.

Li, S. ja Jain, A. (2011). Handbook of Face Recognition. Springer Science & Business Media. 2nd ed.

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation (2018). Future of Humanity Institute, University of Oxford, Centre for the Study of Existential Risk, University of Cambridge, Center for a New American Security, Electronic Frontier Foundation, OpenAI.

Mannila, H. (1996). Data mining: machine learning, statistics, and databases. Int'l Conf. Scientific and Statistical Database Management. IEEE Computer Society.

Manning, C. ja Schütze, H. (1999). Foundations of Statistical Natural Language Processing. MIT Press.

MathWorks (2018). What Is Machine Learning? https://www.mathworks.com/discovery/machine-learning.html (haettu 31.5.2018).

Metz, C. (2018). As China Marches Forward on A.I., the White House Is Silent. https://www.nytimes.com/2018/02/12/technology/china-trump-artificial-intelligence.html (haettu 7.5.2018).

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. ja Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

Muehlhauser, L. ja Salamon, A. (2012). Intelligence Explosion: Evidence and Import. Teoksessa Eden, A., Moor, J., Søraker, J. ja Steinhart, E. (eds.) Singularity Hypotheses. The Frontiers Collection. Springer, Berlin, Heidelberg.

Nevatia, R. (1982). Machine Perception. Prentice-Hall.

OECD (2011). Regulatory reform and innovation. https://www.oecd.org/sti/inno/2102514.pdf (haettu 10.5.2018).

OpenAI, www.openai.com

Otterlo van, M. ja Wiering, M. (2012). Reinforcement learning and Markov decision processes. Reinforcement Learning. Springer Berlin Heidelberg: 3–42.

Picard, R. (1995). Affective Computing. MIT Technical Report 321.

Pichler, M., Bodenhofer, U. ja Schwinger, W. (2004). Context-awareness and artificial intelligence. ÖGAI Journal, 23, 4–11.

Poel van de, I. ja Royakkers, L. (2011). Ethics, Technology, and Engineering – An Introduction. Wiley-Blackwell, Chester, UK.

Polonski, V. (2017). Can we teach morality to machines? Three perspectives on ethics for artificial intelligence. https://www.oii.ox.ac.uk/blog/can-we-teach-morality-to-machines-three-perspectives-on-ethics-for-artificial-intelligence (haettu 9.5.2018).

Ransbotham, S., Kiron, D., Gerbert, P. ja Reeves, M. (2017). Reshaping Business With Artificial Intelligence. MIT Sloan Management Review and The Boston Consulting Group, September 2017.

Russell, S. ja Norvig, P. (2014). Artificial intelligence – A Modern Approach. Prentice Hall.

Sage, A. (1992). Systems Engineering. Wiley IEEE.

Salinas, E. (2018). Katie Malone on Machine Learning. IEEE Computing edge, Feb 2018, pp. 36–38.

Searle, J. (2001). Encyclopedia of cognitive science. Wiley Online Library.

Shapiro, L. ja Stockman, G. (2001). Computer Vision. Prentice Hall.

Sharples, M. (1996). Artificial Intelligence, Human-Computer Interaction, Chapter 10, pp. 293–323. Academic Press.

Sheridan, T. (2016). Human–Robot Interaction, Status and Challenges. Human Factors, June, Volume 58, issue 4, pp. 525–532.

Smalley, E. (2017). AI-powered drug discovery captures pharma interest. Nat Biotechnol. July 12, 35(7), pp. 604–605.

Sommerville, I. (1993). Artificial Intelligence and Systems Engineering, teoksessa Propects of Artificial Intelligence, Sloman, A. (Eds), Hogg, D. (Eds.), Humphreys, G. (Eds.), Ramsay, A. (Eds), Partridge, D. (Eds.), pp. 48–60.

Suksi, M. (2017). On the Openness of the Digital Society: from Religion via Language to Algorithm as the Basis for the Exercise of Public Powers, teoksessa Transparency in the Future – Swedish Openness 250 Years, (eds.) Lind, A. S., Reichel J. ja Österdahl, I. pp. 285–317, Ragulka.

Sun, K., Zhang, J., Zhang, C. ja Hu, J. (2017). Generalized extreme learning machine autoencoder and a new deep neural network. Neurocomputing, 230, 374–381.

Thagard, P. (2008). Cognitive Science, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Zalta. E. (ed.). https://plato.stanford.edu/archives/fall2008/entries/cognitive-science (haettu 11.5.2018).

Tukey, J. (1962). The Future of Data Analysis. Ann. Math. Statist. 33(1962), no. 1, 1–67.

USNWC (2018). Ethics and Emerging Military Technology (EEMT): Home. U.S. Naval War College. https://usnwc.libguides.com/ethics_emerging_military_technology (haettu 11.5.2018).

Vernon, D. (2014). Artificial Cognitive Systems – A Primer. MIT Press.

Vinge, V. (1993). The coming technological singularity: how to survive in the post-human era, teoksessa Proc. Vision 21: interdisciplinary science and engineering in the era of cyberspace, pp. 11–22. NASA: Lewis Research Center.

Vukobratovic, M., Surdilovic, D. ja Ekalo, Y. (2009). Dynamics and Robust Control of Robot – Environment Interaction. New froniters in robotics, Vol 2. World Scientific. New Jersey, USA.

WEF (2016). Top 9 ethical issues in artificial intelligence. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence (haettu 9.4.2018).

Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M., Stolcke, A., Yu D. ja Zweig, G. (2016). Achieving Human Parity in Conversational Speech Recognition. https://arxiv.org/abs/1610.05256

Zhuang, H. ja Roth, Z. (1996). Camera-Aided Robot Calibration. CRC Press. New York, USA.



VALTIONEUVOSTON
SELVITYS- JA TUTKIMUSTOIMINTA

tietokayttoon.fi

ISSN 2342-6799 (pdf)
ISBN 978-952-287-549-5 (pdf)